DOCUMENT RESUME

ED 385 584 TM 024 022

AUTHOR	Henning, Grant; Cascallar, Eduardo
TITLE	A Preliminary Study of the Nature of Communicative Competence.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETS-RR-92-17; TOEFL-RR-36
PUB DATE	Feb 92
NOTE	67p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Adult Students; *Communicative Competence (Languages); *Construct Validity; Definitions; *English (Second Language); *Language Tests; Linguistic Competence; Models; Second Language Instruction; Second Language Learning
IDENTIFIERS	*Test of English as a Foreign Language; Test of Spoken English; Test of Written English

ABSTRACT

Although programs for teaching and testing of English as a second language (ESL) are often evaluated for their attention to communicative competence, there is no consensus on the nature of the communicative competence construct(s) and little empirical support for the various explanatory models. The theoretical literature related to communicative competence was surveyed to identify major variables said to comprise the construct. The presence and measurability of such variables were tested in typical native/nonnative speaker university academic communication, and a tentative model of communicative competence was developed. The relationship of scores from the Test of English as a Foreign Language (TOEFL), the Test of Spoken English (TSE), and the Test of Written English (TWE) to the model was examined with 79 adult ESL students through video recordings and written responses. Results provide information about the comparative contributions of some theory-based communicative competence variables to linguistic, discourse, sociolinguistic, and strategic competencies. Twelve oral and 12 written communication tasks were analyzed and rank ordered for suitability in eliciting communicative language performance. Sixteen tables present analysis results. An appendix lists the ordered tasks. (Contains 28 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 385 584

TOEFL

TEST OF ENGLISH AS A FOREIGN LANGUAGE

Research Reports

REPORT 36
FEBRUARY 1992

A Preliminary Study of the Nature of Communicative Competence

Grant Henning
Eduardo Cascallar

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
H. I. BRAUN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EDUCATIONAL TESTING SERVICE

BEST COPY AVAILABLE

TM024022

A Preliminary Study of the Nature of Communicative Competence

Grant Henning
and
Eduardo Cascallar

Educational Testing Service
Princeton, New Jersey

RR-92-17

Educational Testing Service is an Equal Opportunity/Affirmative Action Employer.

Copyright © 1992 by Educational Testing Service. All rights reserved.

No part of this report may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher. Violators will be prosecuted in accordance with both US and international copyright laws.

EDUCATIONAL TESTING SERVICE, ETS, the ETS logo, TOEFL, the TOEFL logo, TSE, the TSE logo, and TWE are registered trademarks of Educational Testing Service. The TWE logo is a trademark of Educational Testing Service.

The Test of English as a Foreign Language (TOEFL®) was developed in 1963 by a National Council on the Testing of English as a Foreign Language, which was formed through the cooperative effort of more than thirty organizations, public and private, that were concerned with testing the English proficiency of nonnative speakers of the language applying for admission to institutions in the United States. In 1965, Educational Testing Service (ETS) and the College Board assumed joint responsibility for the program, and in 1973, a cooperative arrangement for the operation of the program was entered into by ETS, the College Board, and the Graduate Record Examinations (GRE) Board. The membership of the College Board is composed of schools, colleges, school systems, and educational associations; GRE Board members are associated with graduate education.

ETS administers the TOEFL program under the general direction of a Policy Council that was established by, and is affiliated with, the sponsoring organizations. Members of the Policy Council represent the College Board and the GRE Board and such institutions and agencies as graduate schools of business, junior and community colleges, nonprofit educational exchange agencies, and agencies of the United States government.

A continuing program of research related to the TOEFL test is carried out under the direction of the TOEFL Research Committee. Its six members include representatives of the Policy Council, the TOEFL Committee of Examiners, and distinguished English as a second language specialists from the academic community. Currently the Committee meets twice yearly to review and approve proposals for test-related research and to set guidelines for the entire scope of the TOEFL research program. Members of the Research Committee serve three-year terms at the invitation of the Policy Council; the chair of the committee serves on the Policy Council.

Because the studies are specific to the test and the testing program, most of the actual research is conducted by ETS staff rather than by outside researchers. However, many projects require the cooperation of other institutions, particularly those with programs in the teaching of English as a foreign or second language. Representatives of such programs who are interested in participating in or conducting TOEFL-related research are invited to contact the TOEFL program office. All TOEFL research projects must undergo appropriate ETS review to ascertain that the confidentiality of data will be protected.

Current (1991-92) members of the TOEFL Research Committee are:

James Dean Brown	University of Hawaii
Patricia Dunkel (Chair)	Pennsylvania State University
William Grabe	Northern Arizona University
Kyle Perkins	Southern Illinois University at Carbondale
Elizabeth C. Traugott	Stanford University
John Upshur	Concordia University

Abstract

Since the introduction of the expression "communicative competence" by Hymes in 1972, programs for the teaching and testing of English as a second or foreign language have increasingly been evaluated in terms either of their ability to promote communicative competence or of their sensitivity to a communicative view of language competence and performance. Unfortunately, despite this evident trend, there is no consensus on the nature of the communicative competence construct(s), nor has sufficient empirical support been provided for the various explanatory models proposed.

The present study was conducted to survey the theoretical literature related to communicative competence; to identify major variables said to comprise the construct(s); to test for the comparative presence and measurability of such variables in typical native/nonnative speaker university academic communication; to propose a tentative model of communicative competence as a synthesis of these variables; and to examine the relationship of Test of English as a Foreign Language (TOEFL), Test of Spoken English (TSE), and Test of Written English (TWE) scores with the various elements of the tentative model. Accordingly, a schema was devised for the systematic elicitation of academic communicative language performance and this schema was applied in the assessment of the 79 adult, English-as-a-second language students who were chosen for the study. In all, 158 hours of video-recorded language-interaction episodes and 12 communicative writing samples were gathered by trained interlocutors, and observations on 40 communicative competence variables (21 in an oral modality, and 19 in a written modality, of which 18 in each modality are reported here) were rated by trained evaluators. Variables were examined for reliability and correlation with subtest and total scores of the TOEFL, TSE, and TWE examinations.

Results provide information about the comparative contributions of some theory-based communicative competence variables to domains of linguistic, discourse, sociolinguistic, and strategic competencies. In turn, these competency domains were investigated for their relation to components of language proficiency as assessed by the TOEFL, TSE, and TWE tests. Twelve oral and twelve written communication tasks were also analyzed and rank ordered for suitability in eliciting communicative language performance.

Acknowledgments

Special acknowledgment is due to the following for their substantial contribution to the success of this project:

Dr. Eduardo Cascallar of Educational Testing Service, who assisted the principal author in the training of interlocutors, equipment operators, and raters, and who assisted in data collection and aggregation.

Dr. Marjorie C. Walsleben of the UCLA Extension English Language Center, for help in recruitment of subjects and interlocutors for communicative language interactions, and for coordination and scheduling of on-site training and testing activities.

Vicki Alkadri, Robert Baldwin, Saundra Barnes, Janet Cowal, Ruqayyah Curtis, Kyle Duncan, Laura Earles, Janathan Epstein, Steve Gelman, Stephannie Howard, Erica Levy, Lela Lindsey, Jori Martinez, Maureen Mason, Bruce Miller, Michael Silverman, Judith Tanka, Terry Varvel, and Richard Wolfe of the UCLA Extension English Language Center, for service as proctors, interlocutors, and equipment operators on communicative interaction teams.

Deborah De Marco, Robert Traegler, and Thomas Ward, for service as raters of both written and video-recorded communicative language performance of all subjects in the study.

Joanne Farr of Educational Testing Service for typing the statistical tables in this report.

Table of Contents

Problem 1

Background Literature 2

The Nature and Purpose of Psychological Models 4

Purpose and Scope of Work 5

Method
- Sample 6
- Instrumentation and Procedures 7
- Analyses 11

Results
- Test Descriptive Information 11
- Oral Communication Tasks and Variables 11
- Written Communication Tasks and Variables 12
- Effects of Register and Pragmatic Function on Oral Communication Variables 13
- Effects of Register and Pragmatic Function on Written Communication Variables 14
- Interrelationships among Language Test Scores and Oral Communication Variables 15
- Interrelationships among Language Test Scores and Written Communication Variables 16
- Communication Variables as Predictors of Communicative Competence by Communicative Domain 17
- Oral and Written Task Effectiveness in Representing Elements of Communicative Competence 18
- Correlations of Estimated Communicative Competence Variables with TOEFL, TSE, and TWE Tests and Subtests 19

Discussion and Conclusion 20

Tables 25

Appendix 43

References 49

List of Tables

Table 1 Sample Description 27

Table 2 Design Features 28

Table 3 Means, Standard Deviations, and Reliability Estimates for Standardized Language Tests 29

Table 4 Correlations of TOEFL, TSE, and TWE Subtest and Total Scores 30

Table 5 Means, Standard Deviations, and Interrater Reliabilities for 18 Oral Communication Variables on 12 Oral Communication Tasks 31

Table 6 Means, Standard Deviations, and Interrater Reliabilities for 18 Written Communication Variables on 12 Written Communication Tasks 32

Table 7 Successive Analyses of Variance with 18 Oral Communication Variables as Dependent Variables and Two Levels of Register and Six Levels of Pragmatic Function as Repeated Factors and with Task Sequence Covariance Extracted 33

Table 8 Successive Analyses of Variance with 18 Written Communication Variables as Dependent Variables and Two Levels of Register and Six Levels of Pragmatic Function as Repeated Factors 34

Table 9 Pairwise Correlations of TOEFL, TSE, and TWE Subtest and Total Scores with 18 Oral Communication Variables Averaged across 2 Levels of Register and across 6 Pragmatic Functions 35

Table 10 Pairwise Correlations of TOEFL, TSE, and TWE Subtest and Total Scores with 18 Written Communication Variables Averaged across 2 Levels of Register and across 6 Pragmatic Functions 36

Table 11 Multiple Regression Analyses with Oral Communicative Competence as the Dependent Variable and With Oral Communication Variables by Competence Domain as Independent Variables 37

Table 12 Multiple Regression Analyses with Written Communicative Competence as the Dependent Variable and With Written Communication Variables by Competence Domain as Independent Variables 38

List of Tables *(continued)*

Table 13 Oral Task Effectiveness in Representing Elements of Communicative Competence: Domain Estimates Correlated with Rated Oral Communicative Competence 39

Table 14 Written Task Effectiveness in Representing Elements of Communicative Competence: Domain Estimates Correlated with Rated Written Communicative Competence 40

Table 15 Correlations of Estimated Oral Communicative Competence Variables with TOEFL, TSE, and TWE Tests and Subtests for Six Selected Oral Communication Tasks 41

Table 16 Correlations of Estimated Written Communicative Competence Variables with TOEFL, TSE, and TWE Tests and Subtests for Six Selected Written Communication Tasks 42

Problem

Over the past two decades language acquisition theorists and, more particularly, language testing specialists have argued in support of various theories of language proficiency or communicative competence (Bachman & Palmer, 1981; Canale & Swain, 1980; Carroll, 1980; Hymes, 1972; Oller, 1979). It has been maintained on the one hand that no appropriate, valid assessment of language proficiency is possible without an adequately articulated theory of language proficiency that would go beyond circular operational definitions provided in the tests themselves. On the other hand, it has been acknowledged that, even if it were possible to define and articulate the theoretical constructs perfectly, tests devised to measure the constructs would necessarily be imperfect in their construct validity due to the less-than-perfect comprehensiveness and uniqueness of representative test content (Messick, 1987). Such criticisms of lack of adequate theoretical basis have been made of the Test of English as a Foreign Language (TOEFL) and other standardized tests of language proficiency. In particular, the need to demonstrate empirically the degree to which TOEFL and other similar examinations assess communicative competence in its evolving theoretical formulations has been expressed repeatedly (Bachman, 1986; Duran, Canale, Penfield, Stansfield, & Liskin-Gasparro, 1985; Stansfield, 1983). Until now, however, such a desirable demonstration has had to await better articulation of the communicative competence construct(s).

The problem surfacing around such deliberations is that, although several deductive models of communicative competence/performance or language proficiency have been proposed, none of them has had sufficient empirical basis to permit analysis of the interrelationships among, the relative importances of, or the comprehensiveness of the proposed model components. Perhaps even more fundamental has been the concern expressed that current models of language acquisition and language ability often lack prerequisite specification of what it means to know a language, thus complicating evaluation by allowing different criteria to serve as measures of success (Spolsky, 1985, 1989).

Obviously, language testing and university admissions decision making have had to proceed without the benefit of resolution of the debate over the nature of communicative competence. Nevertheless, there is growing recognition, as indicated by research activity like that reported here, that language testing as reflected in examinations such as TOEFL, the Test of Spoken English (TSE), and the Test of Written English (TWE) can benefit from the results of empirical investigation into the nature of communicative competence. Such benefits include (1) the ability to relate components of language examinations to components of tentative models of communicative competence, (2) the eventual ability to identify and evaluate candidate examination components on the basis of their unique

contributions to the communicative competence constructs as they would be defined, and (3) a demonstration of intent progressively to develop language tests according to a more communicative orientation whenever that should prove empirically beneficial and operationally feasible.

Background Literature

Perhaps the earliest clearly articulated model of language proficiency was set forth by Carroll (1961) and Lado (1961) as a skills/components model. By this model language behavior could be viewed as divisible into four skill areas (i.e., listening, speaking, reading, and writing) and several component categories across these skills (i.e., phonology/graphology, lexis, morphology, and syntax). Although this model has considerable utility and intuitive appeal, subsequent theorists have considered it too restrictive and inadequate to reflect the sociolinguistic, pragmatic, strategic, and discourse-level concerns of "communicative language ability" (Bachman, 1990; Canale & Swain, 1980; Duran, Canale, Penfield, Stansfield, & Liskin-Gasparro, 1985; Faerch & Kasper, 1984; Halliday, 1976; Hymes, 1972; Oller, 1986). With the inclusion of previously neglected written and spoken discourse concerns such as extrasentential cohesion and coherence (Hatch & Long, 1980), the skills/components model might be said to include much of the substance or the "what" of communicative language behavior.

The skills/components model fails, however, to take into consideration sociolinguistic concerns of social register such as those reflected in the formality levels suggested by Joos (1967), that is, frozen, formal, consultative, casual, and intimate levels. The language of the lecture or the seminar is probably different from the language of the sporting event or the dormitory. Such concerns of situational appropriacy might be termed the "who," "when," and "where" of language behavior.

The skills/components model also fails to take into consideration pragmatic concerns related to the purpose of the communicative act such as Halliday's (1973, 1976) classification of language functions into categories of ideational, instrumental, regulatory, interactional, imaginative, and heuristic functions. Rhetorical functions such as narration, description, and comparison may sometimes also fit here as purposeful classifiers to reflect the "why" of communication. Presumably, then, communication could be evaluated, as some have proposed, in terms of the degree to which its objectives are realized, rather than merely by how grammatically accurate its forms are.

The skills/components model appears also to fall short of addressing strategic concerns related to the "how" of communication. Included here are concerns of the rate and density of information transfer and the ability to adapt these to the needs of the interlocutor(s). Also considered under this rubric are concerns

related to the planning, rehearsal, and internal evaluation of communicative acts, including the ability to compensate for linguistic deficiencies as through avoidance strategies. Some of these strategic concerns are reflected in more traditional forms of fluency assessment.

It should be noted that, although the literature on communicative language theory regularly cites various elements of the domains described above (i.e., linguistic [including discourse], sociolinguistic, pragmatic, and strategic competencies), there is no consensus on the nature of the organizational interrelationships among or the relative importance of the various hypothesized components. For example, Hymes (1972) argues that communicative competence is independent of linguistic competence, but Canale and Swain (1980) include linguistic competence within communicative competence to make clear the ultimate inseparability of these domains. Duran et al. (1985) envision a separate discourse domain, merge sociolinguistic and pragmatic domains, and do not include a strategic domain. Bachman and Clark (1987), by way of contrast, believe communicative language proficiency to consist of domains of language competence, strategic competence, and psychophysiological skills. Language competence, they say, consists of organizational and pragmatic competence. Organizational competence is said to consist of grammatical (rules of lexis, morphology, and syntax) and textual (cohesion and rhetorical organization) competence. Pragmatic competence is said to consist of illocutionary competence (language functions) and sociolinguistic competence. Strategic competence is said to involve the ability to recognize, assess, infer, and compensate for deficiencies. Psychophysiological skills are said to consist of the modes and channels (visual, neuromuscular, auditory, and articulatory skills) underlying the ability to listen, speak, read, and write language. Such models of communicative language ability thus tend to be generative rather than taxonomic, unlike the skills/components model (Oller, 1986). And yet there remains very little empirical basis for any hierarchical ordering of the constituent components of these models. Or, as Candlin (1986) aptly expressed it, "such elaborations...offer no explanation of which aspects are used in particular circumstances and how they are drawn upon in the process of making meanings" (p. 41).

Several inferences may be drawn from this brief overview of the theoretical literature on communicative language ability. First, there seems to be consensus among experts that the primary skill domains (i.e., aural/oral, reading/writing) and the elements of language (i.e., phonology, orthography, morphology, syntax, and semantics) as set forth in the earlier skills/components model are necessary but not sufficient to encompass what has thus far come to be known as communicative competence. There also appears to be a consensus that among those missing elements should be a recognition of discourse level concerns such as cohesion and coherence; an acknowledgment of sociolinguistic concerns such as social register and cultural appropriacy; an admission of situational purposes and

pragmatic functions of communication; and an awareness of strategic processes such as use of avoidance and planning strategies in communication. There remains considerable disagreement about the relative importance of these elements for specified communication purposes and about how these agreed elements interrelate in communicative performance. Clearly, any discussion of a theory of communicative competence must still be tentative in nature. One is reminded of an observation by Brennan (1991) that, "from one perspective there are really only two types of theories--those that are known to be wrong and those that are yet to be proven wrong" (p. 262).

The Nature and Purpose of Psychological Models

Before it is possible critically to examine or successfully to construct psychological models, including models of communicative competence or performance, it is useful to reflect on the properties and purposes of such models. Models are constructed to simulate some aspect(s) of reality for some utilitarian purpose(s). Therefore, models should be evaluated not only in terms of how nearly they approximate reality, but more importantly in terms of how well they serve the utilitarian purpose(s) for which they were devised.

One helpful example of model purpose is provided in the field of cartography. Various kinds of two-dimensional maps have been devised as aids to navigation. Some maps are useful geographical models for ocean navigation, others for automobile navigation, and still others for wilderness trekking. In spite of their utility, none of those two-dimensional maps provides a completely accurate representation of three-dimensional reality, nor does any one kind of two-dimensional map serve every navigational purpose equally well. So there is need for many kinds of maps or models, not only because of variation in terrain, but also because of variation in the purposes for encountering the terrain in the first place. Often the value of such maps or models lies not in their sophistication and inclusiveness of detail but in their parsimony and exclusion of all but relevant detail.

In the realm of psychometrics, various mathematical or statistical models have been proposed to accommodate reporting, interpretation, validation, and equating of test scores. Examples such as factor models and item-response-theory models are perhaps even further removed from behavioral reality than two-dimensional maps are from geographical reality. These measurement models are themselves imposed on item and test response data that in turn emanate from tests devised and administered for various purposes according to some articulated or implied theory of behavior. At best, such measurement models can only support or disconfirm the modeling that has already occurred with the construction of the test and the selection of the sample of persons to be tested. This implies that every test is

itself a primitive model of a finite domain of human behavior, constructed for some utilitarian purpose. Tests, like maps, are designed, among other reasons, to reveal position in relation to desired location and salient features of surrounding terrain.

Human behavior, especially language and communication behavior, is likely at least as complex as observable geographical features. A full description of that behavior (i.e., performance) and the capacity or potential for that behavior (i.e., competence) would entail agreement on terminology to represent relevant behavioral construct domains and how they interrelate in form and function, symbol and system, particle and process. Any model of such behavior would necessarily be restricted in scope, selective of reality, and would be conditioned in form by the methods and purpose(s) for model construction. In recognition of these limitations, the present study has intentionally been restricted to the consideration of a finite set of potentially measurable constructs that have been proposed as components of models of communicative competence or of communicative language ability in an academic assessment context.

From this general discussion of models and modeling procedure it should be apparent that our view is that there is no one ultimate model of communicative competence. Nor is it expected that any one model will serve all purposes equally well. Just as there are many situations, purposes, and requirements for language communication, so there are room and need for many models of communicative competence. The authors hold no illusions that either the communicative behaviors considered here or the manners in which they were elicited and evaluated represent any ultimate model of reality. However, we are hopeful that modeling activities such as those reported here can nevertheless be useful in the design of appropriate communicative language assessments.

Purpose and Scope of Work

This study was intended to shed new light on the nature of communicative competence through investigation of the most salient components of the communicative models described above. In particular, we hoped to learn how these components were interrelated, how well they might function in an assessment context, and to what extent they may already be present in various kinds of language assessment. To these ends, the communicative performances of English-as-a-second-language learners were sampled broadly across a variety of communicative genres, skills, purposes, and registers that were said to relate to success in a university academic setting. Once such communicative data were elicited and recorded, it was possible to analyze the data and derive a tentative set of measurable components of communicative language ability. It was possible thereafter to examine ways in which these components appeared to interrelate for the particular group of adult English language learners sampled across a variety of language backgrounds and language proficiency levels. It

was then possible to relate measured elements of communicative performance to component scores and total scores of the TOEFL, TSE, and TWE examinations.

Variables of interest in analysis include a representative set of ratable components of competencies in the linguistic, discoursal, sociolinguistic, pragmatic, and strategic domains. It is important to note that selection of the 18 rated communicative performance variables, the 6 pragmatic functions, the 2 levels of social register, and the 2 paired-skill modalities that were eventually chosen for inclusion in this study was based on a considered rationale. First of all, these variables or constructs were chosen to provide comprehensive representation of the salient components of all the models of communicative performance encountered in the literature; thus, no one particular model was the basis of selection, but the attempt was made to incorporate salient components within domains and across models. In the second place, effort was exerted to gather components of communicative competence that could be considered relevant in the assessment of English for academic purposes. Thus, all of the pragmatic functions elicited and analyzed were functions considered important in an academic context. And finally, it was necessary to limit the selection of variables to those for which there was consensus that they were readily describable, measurable, and reportable. Thus, for example, schematic knowledge, neurolinguistic organization, and illocutionary force were not included in the study.

From these comments it should be apparent that this effort was intended to build on existing models of communicative competence. Modeling attempts reported here constitute an empirical amalgamation of the models cited. This study may also be somewhat unique in the communicative competence literature in that it represents an attempt to provide empirical indication of the interrelatedness, construct validity, and variance accountability of a large number of components of communicative competence simultaneously. In this way, we have tried to span the bridge from theory to data.

Method

Sample

The characteristics of the sample of persons who participated as subjects in the study are summarized in Table 1. Subjects were enrolled as students in an intensive program of English language study in Los Angeles. Most of them had recently arrived in the United States from their native countries and were learning English in order to qualify for eventual admission to a university of their choice. The subjects' class sections were selected by systematic stratification across instructional levels so as to approximate roughly the range and distribution of language proficiency exhibited by those who normally respond to the TOEFL. Assignment of a

communicative interlocutor team to subjects was random. Placement within instructional levels had been determined on the basis of scores obtained on locally developed tests of English language proficiency, and not by use of any of the testing instruments employed in the study. Because of the large number of language tests and oral and written activities each subject was asked to participate in as part of this study, it was recognized from the outset that the sample could not be large. The project proposal called for a sample of 60 persons, but by oversampling we were eventually able to gather data from 79 persons. Subjects participated by informed consent and were paid for their participation at the conclusion of their testing sessions.

As indicated in Table 2, the communicative performance sample consisted of one hour of videotaped oral/aural communicative activities varied in 12 five-minute intervals, and one hour of reading/writing communicative activities also varied in 12 five-minute intervals. Thus, in all there were 12 x 2 x 79 or 1,896 ratable five-minute intervals of communicative performance. These performance intervals were each rated by multiple independent raters on 21 communication variables in one skill modality and 19 communication variables in the other skill modality. Because two of the rated variables in one modality and one of the rated variables in the other modality related to the performance of the native-speaker interlocutor or the naturalness of the encounter rather than to the performance of the subject, only 18 variables in each modality are described in this report. Pilot analyses suggested that use of one-minute intervals as the unit of analysis instead of five-minute intervals, that held promise for increasing the number of observations on each variable by a factor of five, did not provide enough context for reliable assessment.

Instrumentation and Procedures

Standardized Tests. All participant subjects were administered a single recently disclosed version of the Test of English as a Foreign Language (TOEFL), the Test of Spoken English (TSE), and the Test of Written English (TWE), which there was reason to believe none of the subjects had previously encountered. Administration and scoring of these tests conformed to the official methods and procedures, including the time allowed to respond. The TOEFL answer sheets were optically scanned and machine scored. Both the TSE and TWE tests were scored by two officially qualified independent raters, and an external adjudication process was followed in those few cases where ratings were sufficiently discrepant. The TSE examination was recorded using high quality tape recording equipment. To minimize the possibility of subject fatigue, the testing of each subject required two separate days. For about half of the subjects, the TOEFL and TWE tests were administered on the first day, followed by the TSE test and the communicative elicitation activity on the second day. For the other half, the sequence of days was reversed. For no subject did the interval between the two days of testing exceed 10 days. This shortened interval was intended to minimize the possibility that

significant learning or forgetting of English language would take place between test administrations. Table 3 reports the means, standard deviations, and reliability estimates for these examinations as applied to the present sample. Table 4 reports the attenuated (i.e., observed and unadjusted) intercorrelations of all tests and subtests. Note that test and subtest estimates conformed roughly to expectation. Observed low reliability estimates and subsequent intercorrelations for TOEFL subtests were expected due to the smaller numbers of test items within subtests by comparison with the number of items within total tests. Demographic information, including data reported in Table 1, was elicited as part of the standardized testing procedure.

Communication Activities. The schema for elicitation of communicative language activities is reported in Table 2 and elaborated on in the appendix. Note that the schema calls for a succession of 12 five-minute aural/oral activities, and 12 five-minute reading/writing activities. Exactly half of these activities in each modality were designed to reflect informal social register, and the other half were intended to involve more nearly formal social register. Note also that six pragmatic academic functions were identified for language elicitation in both registers and in both modalities. The pragmatic functions of responding to requests for academic information, requesting academic information, persuading change, complaining, apologizing for inadequate behavior, and summarizing information received were determined, after planning and piloting, to represent sufficiently relevant but varied functions for purposes of language assessment. The elicitation schema was also consistent with several of the concerns raised in the theoretical literature on communicative language ability. A function of expressing gratitude was also piloted, but it did not succeed in eliciting sufficient language output for research purposes, and was abandoned. In the case of the aural/oral activities, the two interlocutors were allowed to choose either the formal or informal role according to their own personality preferences, but once the role was decided, they were required to assume that role throughout the interactions. Altogether, 18 persons served on two-person teams as interlocutors and operators of recording equipment. All communication activities, regardless of register, function, or modality, were videotaped for subsequent rating and analysis. In all, 79 subjects were videotaped for two hours each, for a total of 158 hours of recorded communication activity. In addition, for the reading/writing modality, all written material produced was preserved for subsequent scoring and analysis.

Training. Interlocutors were trained by the researchers over a half-day session on two separate days approximately one week prior to the beginning of elicitation. Training included precept, example, and practice opportunities for both language elicitation and use of recording equipment. In each testing encounter, when one of the interlocutors was engaged in language interaction with the subject, the other interlocutor assumed responsibility for recording. Separate

one-day training and practice sessions were also held for the three raters who were prepared to rate independently the oral and written language output on the communicative variables identified. Ratings were checked against criteria outlined in a rating schedule, and corrective feedback was provided before official rating was begun. Both raters and interlocutors were about evenly divided between males and females. Interlocutors were experienced teachers of English as a second language with graduate-level education in applied linguistics or a related field of study.

Rating. Following training and qualifying activities, raters were called upon to make independent ratings of the communicative quality of the oral and written language output of each subject for each five-minute activity. To minimize dependencies among rated variables, raters were required to view and rate each five-minute segment of elicited language across all subjects before moving on to the next five-minute segment. Although this procedure required much more time than would have been necessary had judgments been made across all activities for a given subject in one viewing, it was intended that it would ensure greater independence of the ratings of each successive task. For each communication task, ratings were given on a scale of 0 to 5, where 0 represented a decision that insufficient ratable language data were present to form a judgment on a given communication variable. Thereafter, zeros were treated as instances of missing data, and they tended to occur in just under 5% of all cases. Ratings of 1 to 5 represented judgments as follows: 1, prevents communication, at a novice level; 2, communication with great strain, at an intermediate level; 3, communication with some strain, at an advanced level; 4, communication with only occasional strain, at a superior level; 5, communication with no strain, at the distinguished level of an idealized educated native speaker. In every case where ratings exhibited discrepancies of more than one point, an adjudication process was employed. In the limited cases where adjudication was necessary, it was most frequently achieved through augmenting by one point the rating of the stricter of the two raters. However, in a few cases where the discrepancy exceeded two points, it was also judged necessary to reduce the rating of the more lenient rater by one point.

Communication Variables. After survey of the literature, consideration of design features within the limited time available for rating, and piloting of candidate variables, a series of 18 variables were identified for rating and reporting within the aural/oral language modality (O), and 18 variables were identified within the reading/writing modality (W). These 18 variables within each language modality were purposely as similar as possible across language modalities in order to enable cross-modality generalizations. Variables rated extended across communicative competence domains as indicated in the following schema (Tables 5 through 15 of the results section employ the abbreviated descriptors in parentheses when referring to these variables for successive analyses).

Grammatico-Linguistic Competence

Accuracy of pronunciation (articulation, stress, intonation)	(O1)
Accuracy of orthography (letter formation)	(W1)
Semantic appropriacy of lexis (words, phrases, idioms)	(O2, W2)
Range and sophistication of vocabulary	(O3, W3)
Accuracy of morphology (derivations, affixes)	(O4, W4)
Syntactic accuracy (word order, agreement)	(O5, W5)
Syntactic complexity (length, phrasal/clausal embeddings)	(O6, W6)
Spelling in written communication	(W13)
Punctuation in written communication	(W14)

Discoursal Competence

Discourse cohesion (transitions, conjunctions, continuity)	(O7,W7)
Discourse coherence (pronominal reference, consistency)	(O8, W8)

Sociolinguistic Competence

Appropriateness of register (levels of formality--partially reflected by elicitation context)	(A, B, O9, W9)
Culturally appropriate performance (politeness, naturalness, status-role appropriacy)	(O10, W10)

Pragmatic Competence (reflected by elicitation task as detailed in the appendix)

Responding to requests for academic information	(1)
Requesting academic information	(2)
Persuading change in an academic context	(3)
Complaining in an academic context	(4)
Apologizing for inadequate or offensive behavior	(5)
Summarizing information received	(6)

Strategic Competence

Fluency of oral communication (rate of production)	(O11)
Fluency of written communication (rate of production)	(W11)
Density of information transfer	(O12, W12)
Presence of hesitation phenomena in speaking	(O13)
Nonverbal deficiency compensation in oral communication (gesturing)	(O14)
Verbal deficiency compensation (avoidance, synonomy)	(O15, W15)
Overall strategic success	(O17, W17)
Confidence in oral communication	(O18)
Neatness in written communication	(W18)

In addition, raters were asked to provide judgments of listening (O16) or reading (W16) comprehension as manifested in each communication activity. Most of the variables within domains given

above were amenable to rating within and across both oral and written modalities. Other variables were limited to one or the other modality only, as indicated.

Analyses

Means, standard deviations, and reliability estimates were obtained for all tests, subtests, and communication variables in pertinent contexts. Intercorrelations were computed among all tests, subtests, and communication variables in pertinent contexts. Repeated measures factorial analyses of variance (ANOVAs) were conducted across all ratings within language modalities to test the effects of register, pragmatic function, and the interaction of register and pragmatic function. Multiple correlation and regression analyses were made within communicative competence domains by language modality in order to provide meaningful aggregations of communication variables as preliminary synthetic models. Multiple regression estimates of communicative competence by language modality within competence domain were correlated with actual composite ratings of communicative competence within task in order to determine which pragmatic functions were most effective for assessment. Correlations were disattenuated to control for differences in instrumentation reliability that were due in part to differing numbers of items in tests and subtests. For the repeated measures factorial ANOVAs and again for the multiple correlation and regression analyses, means were imputed for missing data in those few cases where data were incomplete.

Results

1. Test Descriptive Information--Tables 3 and 4

As mentioned in the description of the instrumentation, Tables 3 and 4 report the means, standard deviations, reliability estimates, and intercorrelations for the TOEFL, TSE, and TWE tests, and their respective subtests for the particular sample of persons in this study. As expected, reliability estimates were positively related to test length, so that increasing numbers of like items would be expected to increase reliability estimates within subtests. Subsequent increases in reliability magnitudes would also be expected to increase magnitudes of intercorrelations. These comments are added by way of cautioning absolute comparisons among subtests in Tables 3 and 4. Because of the known relationship between test length and reliability estimation, disattenuated correlations were reported for subsequent analyses (Tables 9, 10, 15, and 16).

2. Oral Communication Tasks and Variables--Table 5

Table 5 reports the means, standard deviations, and interrater

reliabilities observed for 18 communication variables and 12 elicitation tasks in the aural/oral modality. Reliability estimates partially reflect the measurability of the underlying constructs in the contexts and by the methods used for these subjects. By this reliability standard, it can be observed that informal requesting (2A) tended to provide the least successful elicitation context (mean reliability 0.53), and confident speaking (O18) was perhaps the least reliably assessed communication variable (mean reliability 0.66). By way of contrast, responding to informal requests for academic information (1A), summarizing information received (6A and 6B), and use of formal complaints (4B) and apologies (5B) appeared to comprise highly reliable elicitation tasks (mean reliabilities ranging from 0.81 to 0.91). Also, grammatico-linguistic variables (O1-O8) appeared highly measurable overall, as were the strategic variables of rate of production (O11) and information density (O12) by the same criteria. Interestingly, means, standard deviations, and reliability estimates varied widely across communication tasks in accordance with social register or pragmatic function required. This appears to provide preliminary support for a view that communicative language ability is situation specific in that it depends largely on the particular communicative context.

3. Written Communication Tasks and Variables--Table 6

Table 6 reports the means, standard deviations, and interrater reliabilities observed for 18 communication variables in the written modality. Reliabilities below 0.10 are reported as not significant. The lower levels of reliability for ratings of written as opposed to oral performance were due not so much to discrepancies among the ratings assigned by the raters as to the existence of very little variance in the rated variables across subjects. Thus, raters tended to assign 3s and 4s throughout. This may also have been partially due to the relative paucity of written language generated by comparison with the amount of oral language produced. Use of the Spearman-Brown Prophecy Formula to estimate mean test length required to obtain reliabilities comparable to those observed for oral tasks suggested that written tasks appear to require intervals approximately three times longer than oral intervals to permit generation of the same amount of ratable language.

Formal complaining (4B) and informal apologizing (5A) appeared to provide the least satisfactory written elicitation contexts from the viewpoint of allowing reliable and discriminating judgments of performance across subjects. In contrast, ratings of informal (1A) and formal (1B) written responding to requests for academic information showed higher, but still less than desirable, mean reliability estimates (0.59 and 0.64, respectively). Among the written communication variables (W1-W18), judgments of spelling (W13) and neatness (W18) showed the highest mean estimates of reliability (0.73 and 0.82, respectively). The overall means of ratings of written communication reported in Table 6 appeared to be higher and

less variable across levels of formality than the means for oral communication reported in Table 5. These higher means may suggest the possibility that the lower mean reliabilities for ratings of written communicative competence may be partially attributable to a slight ceiling effect; however, the fact that the highest mean reliability estimates were associated with those communication variables with the highest overall means would seem to argue against this possibility. Statistical significance of differences among means by pragmatic function and social register is reported in Tables 7 and 8 and is interpreted below.

4. Effects of Social Register and Pragmatic Function on Oral Communication Variables--Table 7

Successive factorial analyses of variance with repeated measures were conducted to test the significance of mean differences and interactions for the 18 oral communication variables at two levels of social register and in the six pragmatic functions investigated. Mean squares and F-values are reported in Table 7. Due to the presence of measurable order effects for oral communication as elicited, the values reported in Table 7 and the interpretations offered here relate to outcomes after task-sequence covariance has been extracted.

In the case of every rated oral communication variable, level of social register introduced a significant difference in mean rating of communicative performance (i.e., $p < 0.01$ or $p < 0.05$). For oral communication according to every variable tested, this difference was in the direction that formal language communication was uniformly more difficult than informal language communication. Also in the case of every rated oral communication variable, pragmatic function introduced a significant difference among means ($p < 0.01$), suggesting that ratings of communicative competence are highly dependent upon communication context, including both the purpose and the pragmatic function of the communicative act. The easiest pragmatic function throughout was apologizing (5A and 5B), and this ease was most apparent at the informal level of social register (5A). Conversely, the most difficult pragmatic functions throughout was summarizing (6A and 6B), and this difficulty was most visible at the formal level of social register (6B). The significance findings for register and pragmatic function are tempered somewhat by the finding of significant interactions between social register and pragmatic function; however, since the interactions tended to be ordinal as opposed to disordinal in nature, they did not disqualify the earlier reported significance conclusions within the range of abilities considered in the present study.

Inspection of the F-values in Table 7 reveals that the effects of social register were most pronounced for the oral communication variables of cultural appropriacy (O10), strategic success (O17), and confident speaking (O18), in that order. The effects of pragmatic function were most pronounced for the oral communication variables of

cultural appropriacy (O10), register appropriacy (O9), and nonverbal compensation (O14), in that order. Level of social register appeared to have least effect on syntactic accuracy (O5) and the accurate use of morphology (O4), as rated.

In the course of the repeated measures analyses of variance certain procedural steps were found necessary. First of all, it became necessary to replace missing values with mean values to preserve equality of observations within cells. As indicated earlier, less than 5% of the ratings were missing for any rated variable. In the second place, a small but consistent order effect was detected across oral communication tasks (unlike written communication tasks), such that oral communication tended to become easier as the subjects moved through the 12 communication tasks. This warmup effect appears to be particularly evident in second language oral communication assessment. To control for this order effect, task sequential order, 1-12, was correlated separately with ratings on each communication variable, and regression estimates were made of observed ratings. Then, the repeated measures ANOVA was conducted on the residuals of the observed minus estimated ratings. This procedure is analogous to analysis of covariance. The effect of this procedure by comparison with the repeated measures ANOVA conducted on the uncorrected ratings was minimal. However, because level of social register was ordered from informal to formal throughout, control for order effect tended to increase the statistical significance of social register as a main effect. Given the large number of repeated observations, loss of degrees of freedom for this procedure was inconsequential.

Multiple correlation values were computed for each repeated measures ANOVA to ascertain the proportion of variance in each communication variable explained by main effects and interaction effects. R squared values across 18 communication variables ranged from 0.12 to 0.27, and the R squared associated with order effects was 0.09.

5. Effects of Social Register and Pragmatic Function on Written Communication Variables--Table 8

Table 8 reports the results of the same kinds of analyses in the written modality as Table 7 reported in the oral modality. As in the case of the results with the oral modality, pragmatic function was associated with significant mean differences for all or nearly all communication variables. The pragmatic function of responding to requests for academic information (1A and 1B) tended to be the easiest of the six pragmatic functions investigated (overall mean 3.22), and both requesting academic information (2A and 2B) and summarizing information (6A and 6B) tended to be the more difficult of these pragmatic functions as assessed (overall means 3.11 and 3.10 respectively). As indicated by F-values reported in Table 8, effects of pragmatic function were most pronounced for the strategic written communication variables of strategic success (W17), verbal

compensation (W15), and neatness (W18), in that order.

Social register tended to be a nonsignificant effect, except for the written communication variables of verbal compensation (W15), spelling accuracy (W13), and strategic success (W17), in that order. For those three written communication variables with significant effects attributable to social register, as was the case for oral communication variables, the effects were uniformly in the direction that informal written communication was easier than formal written communication. In the case of verbal compensation (W15), the presence of a significant interaction effect makes generalization across strategic function difficult because effects of social register were highly situation dependent. The presence of significant interactions for numerous communication variables in the written modality, as with the oral modality, underscores the complexity of communicative processes and the difficulty faced in formulating any single model of communicative performance that would generalize over all communicative contexts.

As with oral communication repeated measures ANOVA results reported in Table 7, use was made of multiple correlation analysis to determine the proportion of variance in the dependent variables accounted for by the main effects and interaction effects. R squared values ranged from 0.02 to 0.24 for the written communication ANOVAs. This lower proportion of variance accounted for by comparison with the oral communication ANOVAs is no doubt due in part to the lower observed reliabilities for the ratings of written performance. As was suggested earlier, this may have resulted from the use of relatively short time intervals for the elicitation of written language. In subsequent analyses reported in Tables 12, 15, and 16, this problem is partially overcome by reliance on only the most successful communicative tasks for the elicitation of language.

6. Interrelationships among Language Test Scores and Oral Communication Variables Averaged across Levels of Register and Pragmatic Functions--Table 9

Table 9 reports the intercorrelations of the Test of English as a Foreign Language and its subtests, the Test of Spoken English and its components, and the Test of Written English with each of the 18 oral communication variables separately as averaged across the associated two levels of social register and six levels of pragmatic function considered in this study. Because of the earlier reported significant differences across strategic functions, it was also thought important to examine these correlations separately for each level of register within each strategic function. Presentation of the information in Table 9 separately for each pragmatic function at each level of register in each modality would require 12 additional tables identical in format to Table 9 for oral communication variables and 12 additional tables identical in format to Table 10 for written communication variables. Although that information was generated as a

part of this study, in the interest of conservation of space, only Tables 9 and 10 are reported here as cumulative reflections of the kinds of underlying relationships that exist. As one would expect, the magitudes of coefficients within certain of the pragmatic functions and levels of register were much higher than the averaged values reported in Tables 9 and 10.

For samples of 71-79 persons (i.e., 79 responding to the TOEFL and TSE tests and 71 responding to the TSE test), the critical values for correlation coefficients ($p > 0.05$, one tail) are 0.23-0.22. In Tables 9 and 10 the coefficients reported are averaged across 12 observations each, so one would expect the actual critical value to be much lower, closer to that for a sample of 79 times 12, or 948 persons, which would be less than .09. This suggests that, although the reported coefficients in Tables 9 and 10 are not large, they nevertheless tend to exhibit statistical significance throughout. Similarly, recourse to Hotelling's t statistic to test the significance of differences among correlation coefficients for correlated samples suggests that differences as small as .09 among correlation coefficients for some comparisons in Tables 9 and 10 may be significant. The final two columns of Tables 9 and 10 report the mean correlation across variables and the mean disattenuated correlation. Because the standardized subtests varied widely in length and reliability, it is important to hold unreliability constant through disattenuation in order to permit more useful comparisons among subtests.

Interestingly, in the oral modality, TOEFL structure items exhibited the highest mean disattenuated correlation with rated communicative performance (0.53). It is not consistent with usual thinking about communicative language teaching and assessment to expect a high relationship between traditional measures of language structure and contemporary formulations of communicative performance. The Test of Spoken English appeared to be the composite test most related to oral communicative variables as measured here (0.46). This is to be expected. What is less predictable is that TSE measures of grammar (0.51) and fluency (0.50) tended to be slightly more predictive of oral communicative performance overall than was TSE comprehensibility. Perhaps this is due to the fact that comprehensibility as a construct is more heavily dependent on abilities present simultaneously in both the producers and the receivers of language communication; whereas grammar and fluency are more highly associated with the producers alone.

7. Interrelationships among Language Test Scores and Written Communication Variables Averaged across Levels of Register and Pragmatic Functions--Table 10

Table 10 reports the intercorrelations among standardized tests and communication variables as in the case of Table 9, except that Table 10 is reflective of the written modality. Considering the mean

disattenuated correlations in the extreme right-hand column, we can see that TWE total score exhibited the highest correlation of any examination total score measured (0.45), with communication variables overall, in the written modality. As was the case with Table 9, correlations above 0.22 (or even above 0.09 if repeated measures are admitted) achieved statistical significance throughout ($p < 0.05$; one tail). Surprisingly, the subtests of the TOEFL listening component performed comparatively well as predictors of written communication as measured (0.47-0.49). Also surprising is the comparatively low observed correlation of TOEFL structure and written expression with measures of written communication (0.25-0.29), unlike the situation with structure and written expression and communication variables in the oral modality. One could argue that the fact that correlations between ratings of written communicative competence and indirect writing measures (i.e., structure and written expression multiple-choice subtests) were lower than the correlation between ratings of written communicative competence and a direct writing measure (i.e., TWE) is suggestive of a method effect since both TWE ratings and written communicative competence were holistic in nature. However, this explanation does not account for the earlier mentioned high observed correlation between multiple-choice listening scores and holistic writing scores.

8. Communication Variables as Predictors of Communicative Competence by Communicative Domain--Tables 11 and 12

Tables 11 and 12 report successive multiple correlation and regression analyses to enable the meaningful aggregation of communicative variables within their respective communicative competence domains. Again, as was the case with repeated measures ANOVAs reported in Table 7, analyses involving oral communication variables were done after using regression residuals to control for order effects. The dependent variable for the multiple regression analyses reported in Table 11 was oral communicative competence derived as the unweighted combination of the ratings of oral strategic success (O17) and listening comprehension (O16). Similarly, the dependent variable in Table 12 was written communicative competence obtained as the unweighted combination of the ratings of written strategic success (W17) and reading comprehension (W16). In this way, competence within language modality was derived as an independent measure of the ability both to produce and receive language successfully for communicative purposes. The weighted combinations of variables within each competence domain constitute preliminary tentative representations of empirically based modeling attempts.

Note from Table 11 that lexical complexity (O3) and syntactic complexity (O6) were the two communication variables within the linguistic domain contributing most to the prediction of oral communicative competence as measured. Thus, range and sophistication of vocabulary and sentence length along with use of phrasal and clausal embeddings were most predictive of oral communicative

competence in the linguistic domain. Discourse cohesion (O7) (i.e., use of appropriate transitions, conjunctions, and continuity devices) contributed slightly more to the prediction of oral communicative competence within the discourse domain than did discourse coherence. Rated register appropriacy (O9) across informal and formal elicitation contexts was the strongest predictor of oral communicative competence in the sociolinguistic domain. The comparative presence or absence of hesitation phenomena (O13) along with density of information transfer (O12) were the strategic variables most predictive of oral communicative competence as measured. Of the competence domains considered, strategic competence accounted for the most variance in the overall communicative competence construct as measured (i.e., adjusted R squared, 0.866), suggesting the comparative importance of the strategic domain in the assessment of oral communicative competence.

Many of the same comparative contributions found for oral communication variables in Table 11 were repeated for written communication variables in Table 12, with the exceptions that cultural appropriacy (W10) appeared more important than register appropriacy (W9) in the written modality as a predictor of communicative competence in a sociolinguistic domain, and appropriate use of morphology (W4) appeared slightly more predictive of communicative competence as measured than did syntactic complexity (W6). Also, verbal compensation (W15) (i.e., use of avoidance and synonomy strategies) appeared to be the best predictor of written communicative competence as measured within the strategic domain. Once again, the strategic domain appeared most predictive of the overall written communicative competence domain as measured (i.e., selected R squared, 0.490).

In Table 12, adjusted squared multiple correlation coefficients (i.e., selected R squared) are also reported for a selected subsample of the six most successful elicitation tasks as explained in the interpretation of Tables 13 and 14. These additional coefficients are presented in the extreme right-hand column of Table 12. The evident improvement in the values of selected multiple R squared over adjusted multiple R squared values was no doubt due to the elimination of ratings from the six least successful elicitation tasks. This outcome underscores the importance of careful selection of elicitation tasks for the assessment of communicative language abilities.

9. Oral and Written Task Effectiveness in Representing Elements of Communicative Competence - Tables 13 and 14

As noted earlier, wide differences were observed in rated communicative performance that were functions of the particular communicative task employed for the elicitation of language. Tables 5 and 6 reported the comparative difficulties and reliabilities of rated variables by task. A more important way to determine the

comparative success of elicitation tasks is provided through a task by task examination of the magnitudes of correlations between domain estimates of communicative competence and actual independent ratings of communicative competence within language modality. These correlations serve to reflect the comparative validities of elicitation tasks as bases for derivation of communicative competence estimates. Due to the large number of observations reported, resort to Hotellings' t again indicates that differences among correlations as small as .09 may achieve significance ($p < 0.05$, one tail) for some comparisons in Tables 13 and 14. In the extreme right-hand column, tasks are rank ordered for magnitude of mean correlation reported in the preceding column. By this criterion, the six most successful oral elicitation tasks in Table 13 were determined as follows: 1, formal summarizing; 2, informal apologizing; 3, formal apologizing; 4, formal persuading; 5, formal requesting; and 6, informal persuading. Similarly, the six most successful written elecitation tasks in Table 14 were determined as follows: 1, informal responding to requests for academic information; 2, informal summarizing; 3, formal responding to requests for academic information; 4, informal complaining; 5, informal requesting; and 6, formal requesting. On the basis of this prioritizing of elicitation tasks, the final correlations of language tests and subtests with communicative competence domain estimates reported in Tables 15 and 16 were restricted to the six most successful elicitation tasks within each language modality. Also, the selected R squared coefficients reported in the right-hand column of Table 12 were obtained with reference to only the six most successful written elicitation tasks (i.e., 1A, 6A, 1B, 4A, 2A, and 2B).

10. Correlations of Estimated Communicative Competence Variables with TOEFL, TSE, and TWE Tests and Subtests--Tables 15 and 16

Tables 15 and 16 report the correlations of regression estimates of linguistic, discoursal, sociolinguistic, and strategic domains of communicative competence and derived ratings of oral and written communicative competence with the TOEFL, TSE, and TWE tests and subtests. Because test and subtest reliabilities vary as a partial function of numbers of items, partially disattenuated correlations are reported in parentheses. Disattenuations were based on reliabilities of the standardized tests and subtests, but not on reliabilities of accumulated ratings of communicative competence. In both Table 15 and Table 16, correlations are based on observations restricted to the six most successful language elicitation tasks as determined through the analyses reported in Tables 13 and 14. Use was made of regression equations derived from analyses reported in Tables 11 and 12 for providing domain estimates as correlates in Tables 15 and 16.

All tabled coefficients in Tables 15 and 16 were significant beyond the $p < 0.05$ (one tail) level. From these tables comparative judgments can be made of the relative contributions of the various tests and subtests considered to the measurement of components of a preliminary model of communicative competence in an academic setting.

A few words of interpretation: it is apparent from the disattenuated coefficients of Table 15 that the Test of Spoken English and its diagnostic scales were the strongest correlates of oral communicative competence as measured. It is also apparent that, while the subtest of TOEFL structure also performed well as a correlate of oral communicative competence, the written expression subtest of the same component was weakest among subtests for the same assessment purpose. It appears from Table 16 that TOEFL listening in its various subtests performed well as a correlate of written communicative competence. Although this could be due in part to the intrusion of written tasks in the listening component (i.e., correct responding to multiple-choice listening items requires reading comprehension of written response options), it seems equally likely that the ability to attend to some kind of internal dialogue is a correlate of communicative competence in writing. Apart from this interesting phenomenon, it is also evident from Table 16 that the Test of Written English performed comparatively well as a correlate of written communicative competence as assessed.

Discussion and Conclusion

The present study represents a systematic endeavor to gather a large body of communicative performance data from a selected sample of learners of English as a second language and to rate those data on a selected group of communication variables according to a predetermined set of performance criteria. Assessments of the level of communicative performance on each of the respective variables in a variety of academic pragmatic functions and levels of social register were in turn related to scores obtained on total and subtest components of the Test of English as a Foreign Language, the Test of Spoken English, and the Test of Written English.

It is hoped that information of two types may have resulted from this effort: general procedural information that may be of interest to applied linguistic researchers concerning potentially fruitful areas for further empirical enquiry, and test-specific information that may be of use to the TOEFL Committee of Examiners and other persons concerned with the development of language tests and the selection of particular language testing formats.

1. General Procedural Information

It was observed that units of language communication behavior of less than five-minute duration were not generally adequate to provide sufficient context for reliable assessment of the kind conducted here. There were exceptions to this observation; but, since the present study represented an effort to gather comparative information for a

number of communication variables across oral and written modalities simultaneously, it was not feasible here to treat the exceptions exceptionally. Therefore, use was made of a five-minute elicitation procedure throughout. Further, although five-minute elicitation intervals provided adequate samples of oral discourse for reliable rating of variables in the aural/oral mode, lower reliabilities estimated for ratings of communicative behavior in the written modality suggest that future studies should allow more time to elicit written data if the goal of the elicitation is to achieve reliability estimates comparable to those for the oral contexts. Application of the Spearman-Brown Prophecy Formula to mean reliability estimates made for both oral and written data suggested that, for ratings of written passages to exhibit reliability comparable to that of the ratings of oral passages obtained in this study, the written passages could require approximately three times the oral elicitation interval--that is, 15 minutes instead of 5 minutes.

Another way in which the oral elicitation appeared to produce results dissimilar to the written elicitation was related to the presence of sequence or order effects that were detected for the oral tasks, but not for the written tasks. Subjects appeared to improve in their oral communicative interactions as they proceeded through the tasks. For this reason, in the present study it was necessary to extract the variance due to order effects for oral tasks in a manner analogous to analysis of covariance. Whether these order effects were due to articulatory warmup effects, anxiety reduction, or some other cause, it is nevertheless important to consider this phenomenon in the process of establishing oral language assessment contexts.

It was also observed that communicative performance is highly situation specific. In both the aural/oral and the reading/writing modalities, significant interactions were found between social register and pragmatic function for communicative performance variables (e.g., Tables 7 and 8). The presence of such strong and consistent interactions makes difficult the establishment of any single model of communicative competence that could be said to be independent of the particular social register or pragmatic function of the communicative interaction. Highly significant differences in mean performance were found for almost every communication variable in both oral and written modalities, depending on the intended pragmatic function of the communicative act. When communication variables were averaged across elicitation contexts (i.e., social registers and pragmatic functions), correlations with standardized test scores were statistically significant but comparatively low (e.g., Tables 9 and 10). These outcomes strongly suggest that any test of communicative language performance had best specify the kinds of communicative performance (i.e., the pragmatic and situational functions intended) in order to assure validity of assessment.

In general, pragmatic function accounted for more variance in communicative performance than did level of social register as elicited and measured, regardless of modality or variable of interest.

This outcome, coupled with the often significant interaction between register and function, appears to underscore the difficulty of deriving any single global model of communicative competence or performance, and implies the need for a variety of models to serve particular pedagogical or assessment purposes.

2. Test Specific Information: Implications for Communicative Language Test Development

The readers are encouraged to examine Tables 9, 10, 15, and 16 to observe particular relationships of interest for particular language test components depending on language modality and pragmatic function. Several noteworthy inferences may be drawn from these tables.

One surprising outcome is that English language structure as assessed in the TOEFL test related comparatively well to communicative performance as assessed at least in the oral modality registers and functions considered in this study. It would appear that the systematic testing of the rules of English grammar and usage is not necessarily antithetical to concern for the assessment of communicative competence.

Another potentially important outcome, this one evident in Tables 11 and 12, was that, among the various oral and written communicative domains, the strategic domain appeared to provide the best comparative estimates of overall communicative competence as measured. Traditionally, for a variety of psychometric and practical reasons, language testing has not usually focused on fluency of cognition or expression or use of compensatory communication strategies in assessing target language abilities. As commitment grows toward more communicatively oriented language assessment and as new developments arise in the technology and psychometrics of assessment, it will likely be important to reconsider the strategic domain as an important focus of communicative language testing.

With regard to the testing of English language speaking ability, it was interesting to observe that the TSE subtests of grammar and fluency appeared to function at least as well in the oral modality as did the reported decision score on comprehensibility. This may be due in part to the fact that comprehensibility is a complex construct, depending as it does on variables simultaneously at play in both the producer and the receiver of language; whereas the other reported scores (i.e., grammar and fluency) depend primarily on characteristics of the language producer. The task-order effects mentioned in the preceding section have implications for test construction as well. Serious attempts to measure oral communicative competence must take into consideration and control for such progressive effects as articulatory warm-up and anxiety reduction in the design of the language elicitation and the comparison of test methods.

Not surprisingly, listening comprehension as assessed by the

TOEFL test related comparatively well to measures of communicative performance in the oral modality (e.g, Table 15). More surprising was the finding that listening comprehension also related comparatively well to such measures in the written modality (e.g., Table 16). Perhaps the ability to write communicatively is partially a function of the writer's ability to attend to an internal dialogue. This implies the usefulness of an agenda for research into the cognitive processes underlying successful writing performance.

Also in the written modality as reported in Table 16, reading comprehension appeared to be the TOEFL nonlistening subtest most related to communicative performance variables overall. However, serendipitously, the Test of Written English tended to surpass reading comprehension as a correlate of the written communicative competence variables. Also, a positive outcome reported in Table 10 was the finding that performance on the Test of Written English was more related to strategic success in the written modality (W17) than was any other test or subtest considered. Parenthetically, this constitutes further evidence of the content and construct validity of the Test of Written English.

Clearly, many additional kinds of analyses should be conducted with the present data set, and the authors are currently planning additional kinds of analyses to test particular hypotheses concerning communicative language ability. As it was, more than 200 pages of statistical tables were generated from the current data set. Far more information was available from the study than could be reported and interpreted in a single report of this kind.

The most important recommendations the authors wish to make from the present study are, first, that any valid assessment of communicative competence can take place only within a well-articulated framework for the elicitation and rating of communicative language abilities. Because communicative performance was found to be highly dependent on particulars of assessment context, communication purpose, and pragmatic function, care must be taken to ensure that appropriate contexts are defined before communicative language tests are devised. A secondary major observation is that the evidence gathered here strongly suggests that traditional measures of knowledge and use of appropriate language structure, as represented in the structure components of the TOEFL and TSE tests, are not empirically unrelated to ratings of communicative competence in any of its linguistic, discoursal, sociolinguistic, or strategic domains. In the legitimate endeavor to develop language proficiency tests that are more communicatively oriented, we should not abandon concern for direct assessment of language structure. And finally, based on the observed importance of the strategic domain in accounting for variance in the communicative competence construct as defined here, the authors would recommend renewed commitment to the measurement of abilities such as fluency of cognition and expression, density of information transfer, and compensatory communication strategies in the assessment of communicative language ability.

TABLES

TABLE 1

SAMPLE DESCRIPTION

Language	N	Age	N	Sex	N
Arabic	1	11-15	1	Female	30
Chinese	10	16-20	13	Male	49
French	3	21-25	40		
German	2	26-30	13		
Indonesian	3	31-35	7		
Italian	6	36-40	3		
Japanese	12	41-45	1		
Korean	21	46-50	1		
Portuguese	5				
Russian	1				
Spanish	9				
Thai	5				
Turkish	1				
Total	79	Total	79	Total	79

TABLE 2

DESIGN FEATURES

STEP 1 - Training of interlocutors and video machine operators (two half-day group training sessions)

STEP 2 - Administration of TOEFL, TWE, and demographic information tests (two one-day group administrations)

STEP 3 - Administration of the Test of Spoken English and the oral and written communication tasks (individual half-day sessions scheduled for each examinee)

Communication tasks consisted of 12, 5-minute videotaped interactive sessions administered separately in both oral and written modalities, as follows:

Task	Time (minutes)	Pragmatic Function	Register
1	5	Responding to requests for academic information	Informal
2	5	Responding to requests for academic information	Formal
3	5	Requesting academic information	Informal
4	5	Requesting academic information	Formal
5	5	Persuading change in an academic context	Informal
6	5	Persuading change in an academic context	Formal
7	5	Complaining in an academic context	Informal
8	5	Complaining in an academic context	Formal
9	5	Apologizing in an academic context	Informal
10	5	Apologizing in an academic context	Formal
11	5	Summarizing academic discourse	Informal
12	5	Summarizing academic discourse	Formal

Exact elicitation prompts are provided in the appendix.

STEP 4 - Scoring tests and rating oral and written behavior according to 18 oral and 18 written communication variables

TABLE 3

MEANS, STANDARD DEVIATIONS, AND RELIABILITY ESTIMATES FOR STANDARDIZED LANGUAGE TESTS

Test	Maximum Possible	Mean	S	R[a]
TOEFL Total (Raw)	146	85.063	18.363	.921
TOEFL Total (Scaled)	677	478.861	51.252	.921
Listening	50	29.734	7.455	.841
Listening A	20	11.962	3.256	.685
Listening B	15	9.127	2.549	.601
Listening C	15	8.646	2.660	.605
Structure & Written Expression	38	22.114	5.484	.763
Structure	14	8.797	2.503	.622
Written Expression	24	13.315	3.750	.665
Vocabulary & Reading	58	33.215	8.150	.842
Vocabulary	29	17.430	3.865	.655
Reading	29	15.785	5.161	.804
TSE Comprehensibility	300	147.288	46.971	.902
Pronunciation	3	1.502	0.485	.860
Grammar	3	1.803	0.467	.810
Fluency	3	1.464	0.492	.837
TWE	6	3.032	0.794	.860

[a] R = Kuder-Richardson Formula 20 for TOEFL (N = 79); mean interskill correlation for TSE (N = 71); interrater reliability for TWE (N = 79).

TABLE 4

CORRELATIONS OF TOEFL, TSE, AND TWE SUBTEST AND TOTAL SCORES
(N = 79 TOEFL, 71 TSE, 71 TWE)

	TT	T1	T1A	T1B	T1C	T2	T2A	T2B	T3	T3A	T3B	CO	PR	GR	FL	TWE
TOEFL Total (TT)	1.00															
Listening (T1)	.87	1.00														
Listening A (T1A)	.82	.92	1.00													
Listening B (T1B)	.65	.86	.70	1.00												
Listening C (T1C)	.79	.85	.66	.57	1.00											
Structure and Written Expression (T2)	.84	.61	.57	.43	.62	1.00										
Structure (T2A)	.68	.55	.49	.44	.51	.81	1.00									
Written Expression (T2B)	.78	.53	.51	.33	.53	.92	.52	1.00								
Vocabulary and Reading (T3)	.90	.63	.62	.40	.60	.66	.49	.64	1.00							
Vocabulary (T3A)	.77	.48	.48	.25	.52	.62	.46	.60	.87	1.00						
Reading (T3B)	.84	.62	.62	.45	.56	.58	.42	.57	.93	.62	1.00					
TSE																
Comprehensibility (CO)	.52	.59	.52	.58	.44	.35	.43	.22	.38	.29	.38	1.00				
Pronunciation (PR)	.49	.57	.46	.57	.47	.36	.50	.20	.32	.23	.34	.94	1.00			
Grammar (GR)	.51	.58	.53	.59	.40	.36	.43	.24	.36	.29	.36	.86	.80	1.00		
Fluency (FL)	.46	.58	.51	.56	.43	.27	.34	.17	.31	.21	.34	.91	.84	.76	1.00	
TWE	.63	.53	.53	.41	.45	.48	.37	.45	.60	.45	.61	.44	.41	.51	.45	1.00

TABLE 5

MEANS, STANDARD DEVIATIONS, AND INTERRATER RELIABILITIES FOR 18 ORAL COMMUNICATION VARIABLES ON 12 ORAL COMMUNICATION TASKS

TASK	N		O1	O2	O3	O4	O5	O6	O7	O8	O9	O10	O11	O12	O13	O14	O15	O16	O17	O18	VARIABLE LABEL
1A	69	M	1.85	1.86	1.96	1.85	1.73	1.85	1.88	2.82	2.25	2.46	2.07	2.01	1.87	2.01	2.01	2.18	2.01	2.89	O1 = Pronunciation
Informal		S	.54	.63	.67	.64	.66	.72	.71	.67	.51	.57	.78	.80	.74	.75	.81	.77	.82	.72	O2 = Lexical Accuracy
Responding		R	.73	.81	.82	.79	.82	.85	.81	.79	.66	.82	.86	.85	.84	.86	.87	.84	.90	.71	O3 = Lexical Complexity
1B	69	M	2.22	2.20	2.29	2.23	2.19	2.25	2.26	2.23	2.38	2.24	2.39	2.44	2.41	2.31	2.27	2.52	2.46	3.34	O4 = Morphology
Formal		S	.51	.55	.55	.46	.56	.54	.50	.50	.50	.48	.70	.72	.64	.50	.50	.66	.77	.70	O5 = Syntactic Accuracy
Responding		R	.77	.78	.78	.61	.80	.81	.71	.66	.77	.77	.90	.91	.79	.73	.74	.86	.91	.68	O6 = Syntactic Complexity
2A	70	M	2.11	2.16	2.16	2.21	2.16	2.18	2.26	2.26	2.25	2.25	2.17	2.16	2.23	2.21	2.21	2.34	2.29	3.30	O7 = Discourse Cohesion
Informal		S	.43	.45	.45	.35	.43	.42	.32	.32	.30	.30	.44	.47	.39	.36	.36	.41	.52	.51	O8 = Discourse Coherence
Requesting		R	.65	.62	.62	.45	.62	.56	.40	.40	.34	.34	.60	.65	.60	.47	.47	.57	.66	.53	O9 = Register Appropriacy
2B	70	M	2.24	2.26	2.26	2.27	2.24	2.24	2.29	2.29	2.29	2.29	2.24	2.23	2.29	2.31	2.31	2.29	2.29	3.34	O10 = Cultural Appropriacy
Formal		S	.37	.37	.37	.36	.47	.47	.40	.40	.40	.38	.47	.49	.39	.41	.43	.51	.52	.59	O11 = Rate of Production
Requesting		R	.56	.45	.45	.50	.66	.66	.59	.55	.56	.56	.66	.68	.52	.51	.55	.71	.69	.65	O12 = Information Density
3A	70	M	2.37	2.38	2.38	2.38	2.38	2.39	2.38	2.38	2.38	2.39	2.41	2.40	2.40	2.39	2.39	2.43	2.47	3.45	O13 = Hesitation Phenomena
Informal		S	.40	.40	.40	.38	.39	.38	.40	.40	.39	.40	.45	.45	.41	.41	.42	.43	.47	.55	O14 = Nonverbal Compensation
Persuading		R	.56	.59	.59	.44	.49	.47	.56	.56	.52	.57	.62	.63	.57	.63	.64	.61	.64	.56	O15 = Verbal Compensation
3B	70	M	2.34	2.34	2.34	2.35	2.34	2.34	2.33	2.33	2.33	2.34	2.33	2.33	2.34	2.34	2.37	2.36	2.38	3.25	O16 = Listening Comprehension
Formal		S	.42	.42	.42	.41	.41	.42	.41	.41	.41	.42	.41	.42	.41	.41	.42	.40	.44	.50	O17 = Strategic Success
Persuading		R	.72	.73	.73	.65	.68	.71	.68	.68	.68	.73	.68	.65	.68	.68	.60	.57	.61	.66	O18 = Confident Speaking
4A	71	M	2.56	2.59	2.57	2.54	2.54	2.55	2.67	2.66	2.67	2.71	2.78	2.73	2.85	2.73	2.70	2.89	2.74	3.64	
Informal		S	.50	.48	.50	.48	.46	.49	.44	.45	.48	.45	.55	.56	.56	.53	.47	.49	.59	.59	
Complaining		R	.83	.74	.75	.71	.68	.76	.65	.70	.73	.72	.74	.71	.65	.64	.64	.67	.71	.64	
4B	71	M	2.55	2.54	2.54	2.52	2.53	2.53	2.54	2.54	2.55	2.54	2.54	2.54	2.50	2.54	2.51	2.52	2.51	3.33	
Formal		S	.49	.51	.51	.50	.51	.51	.52	.52	.49	.51	.51	.53	.53	.51	.51	.53	.54	.67	
Complaining		R	.80	.87	.87	.83	.85	.85	.79	.80	.80	.85	.82	.85	.84	.85	.85	.81	.76	.76	
5A	71	M	2.81	2.82	2.82	2.82	2.83	2.83	2.83	2.82	2.83	2.85	2.84	2.84	2.84	2.84	2.84	2.84	2.83	3.61	
Informal		S	.35	.36	.36	.35	.36	.36	.36	.38	.36	.33	.35	.35	.33	.34	.34	.36	.36	.48	
Apologizing		R	.66	.65	.65	.72	.71	.71	.71	.68	.71	.62	.67	.67	.62	.61	.61	.67	.66	.61	
5B	71	M	2.73	2.73	2.73	2.78	2.76	2.76	2.80	2.79	2.81	2.80	2.80	2.78	2.80	2.80	2.80	2.85	2.78	3.57	
Formal		S	.45	.46	.46	.44	.45	.45	.42	.42	.40	.41	.40	.44	.43	.43	.42	.38	.45	.60	
Apologizing		R	.90	.92	.92	.95	.93	.93	.91	.88	.87	.89	.84	.78	.89	.89	.91	.76	.79	.72	
6A	71	M	2.68	2.68	2.68	2.68	2.69	2.68	2.67	2.70	2.70	2.69	2.67	2.69	2.70	2.70	2.71	2.80	2.79	3.54	
Informal		S	.43	.44	.44	.44	.44	.46	.48	.44	.44	.44	.46	.48	.44	.44	.43	.50	.60	.54	
Summarizing		R	.80	.81	.81	.81	.81	.81	.83	.83	.83	.83	.81	.85	.80	.80	.76	.67	.80	.71	
6B	68	M	2.41	2.40	2.40	2.40	2.40	2.40	2.39	2.39	2.39	2.40	2.38	2.37	2.39	2.38	2.38	2.39	2.37	3.11	
Formal		S	.56	.56	.56	.56	.56	.56	.59	.59	.59	.58	.59	.60	.57	.58	.58	.61	.61	.57	
Summarizing		R	.93	.92	.92	.92	.92	.92	.95	.95	.95	.94	.92	.90	.90	.90	.90	.94	.91	.65	

BEST COPY AVAILABLE

TABLE 6

MEANS, STANDARD DEVIATIONS, AND INTERRATER RELIABILITIES FOR 18 WRITTEN COMMUNICATION VARIABLES ON 12 WRITTEN COMMUNICATION TASKS

TASK	N		W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16	W17	W18	VARIABLE LABEL
1A	76	M	3.06	2.91	2.85	2.82	2.73	2.81	2.90	2.90	2.95	3.09	3.07	2.90	4.41	4.22	3.53	3.49	3.31	3.66	W1 = Orthography
Informal		S	.68	.61	.60	.57	.60	.58	.52	.55	.51	.45	.52	.58	.69	.60	.42	.44	.57	.71	W2 = Lexical Accuracy
Responding		R	.73	.68	.58	.59	.64	.64	.62	.68	.62	.47	.47	.53	.75	.60	.46	.25	.46	.77	W3 = Lexical Complexity
1B	73	M	3.06	2.93	2.95	3.02	2.84	2.95	3.04	3.10	3.11	3.10	3.12	3.02	4.14	4.28	3.18	3.53	3.28	3.45	W4 = Morphology
Formal		S	.51	.58	.58	.47	.58	.54	.51	.48	.46	.44	.46	.57	1.00	.57	.45	.41	.58	.70	W5 = Syntactic Accuracy
Responding		R	.58	.75	.77	.57	.66	.63	.60	.63	.63	.55	.54	.70	.87	.70	.48	.35	.61	.87	W6 = Syntactic Complexity
2A	75	M	2.88	2.75	2.75	2.90	2.67	2.73	2.91	2.93	3.01	3.03	2.98	2.86	4.21	4.13	3.10	3.42	3.19	3.32	W7 = Discourse Cohesion
Informal		S	.30	.39	.38	.34	.43	.42	.34	.31	.23	.24	.35	.38	.94	.59	.31	.42	.52	.66	W8 = Discourse Coherence
Requesting		R	.15	.56	.54	.40	.58	.61	.48	.44	.31	.22	.49	.51	.83	.49	.44	.59	.72	.93	W9 = Register Appropriacy
2B	75	M	2.94	2.72	2.73	2.95	2.64	2.73	2.96	2.95	2.98	2.99	3.01	2.89	4.00	4.29	3.05	3.51	3.09	3.61	W10 = Cultural Appropriacy
Formal		S	.25	.34	.34	.25	.38	.34	.24	.25	.19	.18	.19	.30	.91	.48	.23	.18	.41	.74	W11 = Rate of Production
Requesting		R	N.S.	.33	.41	.39	.46	.33	.20	.17	.42	.45	.38	.33	.81	.62	N.S.	N.S.	.64	.92	W12 = Information Density
3A	73	M	2.96	2.85	2.87	2.94	2.80	2.85	2.99	2.95	3.01	3.03	3.03	2.95	4.30	4.37	3.03	3.39	2.97	3.74	W13 = Spelling
Informal		S	.23	.32	.32	.29	.34	.33	.24	.25	.19	.18	.19	.30	.91	.48	.23	.18	.41	.74	W14 = Punctuation
Persuading		R	.55	.41	.42	.53	.45	.42	.30	.28	.34	N.S.	N.S.	.44	.74	.56	N.S.	.63	.54	.72	W15 = Verbal Compensation
3B	73	M	2.93	2.81	2.80	2.89	2.70	2.80	2.91	2.90	2.93	2.95	2.95	2.87	4.20	4.30	3.06	3.38	2.93	3.78	W16 = Reading Comprehension
Formal		S	.24	.36	.35	.25	.41	.34	.27	.29	.24	.19	.21	.32	.70	.42	.28	.36	.41	.64	W17 = Strategic Success
Persuading		R	.44	.68	.53	.20	.61	.43	.67	.55	.44	N.S.	.33	.51	.75	.70	.16	.37	.57	.91	W18 = Neatness
4A	75	M	2.89	2.83	2.81	2.81	2.71	2.83	2.89	2.89	2.89	2.98	2.93	2.91	4.41	4.40	3.01	3.51	3.04	3.87	
Informal		S	.25	.30	.32	.36	.40	.31	.26	.27	.26	.15	.22	.26	.61	.27	.16	.25	.39	.55	
Complaining		R	N.S.	.36	.48	.58	.64	.50	.37	.35	.37	N.S.	.23	.39	.54	.44	N.S.	N.S.	.44	.82	
4B	74	M	2.89	2.87	2.85	2.79	2.66	2.87	2.94	2.94	2.96	2.99	2.98	2.95	4.33	4.48	3.01	3.36	2.89	4.07	
Formal		S	.32	.29	.30	.36	.42	.28	.23	.23	.22	.19	.19	.23	.59	.31	.18	.40	.35	.55	
Complaining		R	.17	N.S.	N.S.	.40	.44	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	.67	.32	.45	.63	.60	.52	
5A	75	M	3.00	2.91	2.90	2.93	2.81	2.93	3.01	3.01	3.03	3.03	3.02	3.03	4.15	4.40	3.02	3.49	2.97	4.02	
Informal		S	.25	.28	.29	.27	.39	.26	.19	.17	.14	.14	.15	.26	.73	.23	.13	.25	.25	.66	
Apologizing		R	N.S.	.30	.30	.48	.65	.12	.44	N.S.	N.S.	N.S.	N.S.	.68	.66	.53	N.S.	.47	.52	.93	
5B	75	M	2.97	2.91	2.85	2.83	2.67	2.93	2.94	2.95	2.97	2.99	2.95	2.92	4.17	4.31	3.02	3.47	2.94	3.85	
Formal		S	.27	.30	.32	.31	.45	.25	.26	.25	.22	.18	.24	.29	.68	.45	.17	.30	.41	.75	
Apologizing		R	.28	.56	.49	.42	.61	.37	.57	.61	.61	.55	.47	.61	.68	.80	N.S.	.33	.55	.91	
6A	74	M	2.94	2.90	2.80	2.83	2.78	2.87	2.89	2.89	2.90	2.89	2.83	2.81	4.41	4.38	2.93	3.26	2.66	3.79	
Informal		S	.25	.29	.38	.35	.36	.32	.32	.32	.27	.30	.37	.39	.61	.40	.33	.57	.64	.61	
Summarizing		R	.27	.13	.41	.51	.17	.34	.41	.41	N.S.	.40	.51	.51	.71	.19	.44	.69	.66	.86	
6B	55	M	2.96	2.92	2.86	2.94	2.86	2.88	2.93	2.92	2.96	2.96	2.91	2.87	4.18	4.45	2.94	2.95	2.52	3.86	
Formal		S	.29	.23	.31	.22	.30	.27	.22	.23	.20	.20	.24	.29	.59	.21	.22	.45	.47	.63	
Summarizing		R	.40	N.S.	.37	N.S.	.22	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	.70	.47	.32	.39	.52	.69	

Note: N.S. = Reliability <0.10

BEST COPY AVAILABLE

TABLE 7

SUCCESSIVE ANALYSES OF VARIANCE WITH 18 ORAL COMMUNICATION VARIABLES AS DEPENDENT VARIABLES AND TWO LEVELS OF REGISTER AND SIX LEVELS OF PRAGMATIC FUNCTION AS REPEATED FACTORS AND WITH TASK SEQUENCE COVARIANCE EXTRACTED (N = 79)

		01		02		03		04		05		06		07		08		09	
TASK SOURCE	DF	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F
REGISTER	1	.44	4.83**	.86	8.50*	.69	7.47*	.51	4.81**	.43	3.94	.95	9.45*	1.17	11.79*	1.00	9.86*	1.72	17.72*
PRAGMATIC FUNCTION	5	1.85	15.57*	1.88	14.46*	1.65	13.43*	2.04	16.56*	2.26	16.46*	1.96	14.29*	2.34	18.39*	2.39	18.50*	2.34	21.31*
SUBJECTS	78	1.20		1.27		1.35		1.11		1.28		1.31		1.22		1.19		1.10	
REGISTER X PRAGMATIC FUNCTION	5	1.45	16.26*	1.32	14.05*	1.10	11.48*	1.39	15.39*	1.81	18.32*	1.40	14.10*	1.55	17.04*	1.60	17.03*	.67	8.46*
REGISTER X SUBJECTS	78	.09		.09		.10		.11		.11		.10		.10		.10		.10	
PRAGMATIC FUNCTION X SUBJECTS	390	.12		.13		.12		.12		.14		.14		.13		.13		.11	
REGISTER X PRAGMATIC FUNCTION X SUBJECTS	390	.09		.09		.10		.09		.10		.10		.09		.09		.08	

		010		011		012		013		014		015		016		017		018	
TASK SOURCE	DF	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F
REGISTER	1	2.99	31.48*	1.62	16.53*	.98	8.79*	1.81	17.22*	.93	7.65*	1.56	13.12*	3.40	29.34*	2.12	17.41*	2.29	22.47*
PRAGMATIC FUNCTION	5	2.79	26.04*	2.80	19.59*	2.66	17.29*	2.50	18.23*	2.55	20.43*	2.12	15.70*	2.30	16.33*	1.58	10.21*	1.82	10.41*
SUBJECTS	78	1.08		1.66		1.69		1.37		1.26		1.29		1.48		1.90		2.04	
REGISTER X PRAGMATIC FUNCTION	5	.43	4.81**	1.54	16.02*	1.96	17.19*	2.26	20.73*	2.11	19.17*	1.30	12.51*	2.30	19.02*	2.62	18.43*	2.96	20.03*
REGISTER X SUBJECTS	78	.10		.10		.11		.10		.12		.12		.12		.12		.10	
PRAGMATIC FUNCTION X SUBJECTS	390	.11		.14		.15		.14		.13		.14		.14		.16		.18	
REGISTER X PRAGMATIC FUNCTION X SUBJECTS	390	.09		.10		.11		.11				.10		.12		.14		.15	

VARIABLE LABEL

01 = Pronunciation Accuracy
02 = Lexical Accuracy
03 = Lexical Complexity
04 = Morphology
05 = Syntactic Accuracy
06 = Syntactic Complexity
07 = Discourse Cohesion
08 = Discourse Coherence
09 = Register Appropriacy
010 = Cultural Appropriacy
011 = Rate of Production
012 = Information Density
013 = Hesitation Phenomena
014 = Nonverbal Compensation
015 = Verbal Compensation
016 = Listening Comprehension
017 = Strategic Success
018 = Confident Speaking

*p < .01
**p < .05

TABLE 8

SUCCESSIVE ANALYSES OF VARIANCE WITH 18 WRITTEN COMMUNICATION VARIABLES AS DEPENDENT VARIABLES AND TWO LEVELS OF REGISTER AND SIX LEVELS OF PRAGMATIC FUNCTION AS REPEATED FACTORS (N = 79)

		W1		W2		W3		W4		W5		W6		W7		W8		W9	
TASK SOURCE	DF	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F
REGISTER	1	.01	.07	.00	.05	.01	.07	.19	2.37	.14	1.13	.09	.95	.07	.84	.20	2.47	.10	1.58
PRAGMATIC FUNCTION	5	.51	6.13*	.66	6.21*	.40	4.03*	.29	3.25**	.46	4.11*	.64	6.79*	.14	1.99	.19	2.53**	.24	3.69*
SUBJECTS	78	.43		.52		.57		.40		.69		.50		.52		.34		.25	
REGISTER X PRAGMATIC FUNCTION	5	.04	1.31	.03	.34	.13	1.39	.41	4.46*	.27	1.92	.14	1.68	.23	3.30*	.30	4.17*	.27	4.84*
REGISTER X SUBJECTS	78	.07		.09		.12		.08		.12		.10		.08		.08		.06	
PRAGMATIC FUNCTION X SUBJECTS	390	.08		.11		.10		.09		.11		.09		.07		.07		.07	
REGISTER X PRAGMATIC FUNCTION X SUBJECTS	390	.03		.09		.09		.09		.14		.09		.07		.07		.06	

		W10		W11		W12		W13		W14		W15		W16		W17		W18	
TASK SOURCE	DF	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F	MS	F
REGISTER	1	.03	.54	.03	.33	.05	.45	4.18	11.12*	.21	1.20	.70	13.21*	.35	3.08	.92	6.35*	.24	1.37
PRAGMATIC FUNCTION	5	.40	7.76*	.65	9.50*	.32	3.31*	1.32	3.62*	1.09	6.62*	2.93	39.62*	2.30	19.30*	6.53	42.95*	5.79*	25.28*
SUBJECTS	78	.19		.24		.39		1.93		.44		.16		.34		.71		2.39	
REGISTER X PRAGMATIC FUNCTION	5	.11	2.28**	.17	2.52**	.27	2.82**	.40	1.00	.04	.35	.84	16.84*	.43	3.33*	.09	.54	1.35	6.59*
REGISTER X SUBJECTS	78	.06		.08		.10		.38		.17		.05		.11		.14		.18	
PRAGMATIC FUNCTION X SUBJECTS	390	.05		.07		.10		.36		.16		.07		.12		.15		.23	
REGISTER X PRAGMATIC FUNCTION X SUBJECTS	390	.05		.07		.10		.40		.13		.05		.13		.17		.21	

VARIABLE LABEL

W1 = Orthography
W2 = Lexical Accuracy
W3 = Lexical Complexity
W4 = Morphology
W5 = Syntactic Accuracy
W6 = Syntactic Complexity
W7 = Discourse Cohesion
W8 = Discourse Coherence
W9 = Register Appropriacy
W10 = Cultural Appropriacy
W11 = Rate of Production
W12 = Information Density
W13 = Spelling
W14 = Punctuation
W15 = Verbal Compensation
W16 = Reading Comprehension
W17 = Strategic Success
W18 = Neatness

*p < .01
**p < .05

TABLE 9

PAIRWISE CORRELATIONS OF TOEFL, TSE, AND TWE SUBTEST AND TOTAL SCORES
With 18 ORAL COMMUNICATION VARIABLES AVERAGED ACROSS 2 LEVELS OF REGISTER
AND ACROSS 6 PRAGMATIC FUNCTIONS
(N = 79 TOEFL, 71 TSE, 79 TWE)

	01	02	03	04	05	06	07	08	09	010	011	012	013	014	015	016	017	018	R	R^d
TOEFL Total	.35	.35	.43	.36	.33	.32	.42	.33	.36	.30	.44	.42	.41	.34	.31	.44	.40	.28	.37	.43
Listening	.36	.34	.41	.33	.31	.36	.37	.32	.30	.25	.38	.39	.37	.28	.26	.40	.37	.26	.24	.42
Listening A	.29	.27	.40	.29	.23	.31	.34	.25	.29	.26	.36	.37	.33	.25	.22	.40	.34	.27	.30	.41
Listening B	.27	.26	.28	.25	.29	.29	.27	.31	.23	.17	.22	.28	.28	.21	.21	.32	.29	.18	.26	.38
Listening C	.38	.36	.39	.33	.31	.33	.36	.29	.24	.21	.41	.38	.36	.27	.26	.31	.33	.23	.32	.46
Structure and Written Expression	.35	.33	.38	.32	.30	.33	.37	.30	.32	.22	.35	.33	.31	.31	.29	.36	.32	.24	.32	.41
Structure	.41	.39	.44	.33	.35	.36	.39	.36	.35	.32	.39	.36	.35	.39	.37	.36	.38	.34	.37	.53
Written Expression	.23	.22	.25	.24	.20	.24	.28	.19	.23	.11	.25	.23	.21	.17	.18	.28	.20	.12	.21	.29
Vocabulary and Reading	.24	.28	.35	.30	.27	.27	.36	.25	.34	.30	.41	.39	.39	.31	.28	.40	.35	.24	.32	.39
Vocabulary	.22	.19	.27	.23	.18	.18	.27	.15	.25	.22	.36	.30	.33	.24	.20	.29	.28	.20	.24	.33
Reading	.24	.29	.35	.29	.28	.28	.37	.28	.35	.31	.37	.39	.37	.31	.29	.42	.35	.24	.32	.40
TSE																				
Comprehensibility	.40	.37	.42	.40	.44	.38	.36	.41	.38	.30	.44	.43	.40	.40	.36	.43	.40	.34	.39	.46
Pronunciation	.38	.33	.35	.33	.37	.29	.30	.36	.31	.23	.37	.35	.33	.34	.29	.35	.34	.29	.33	.40
Grammar	.37	.39	.46	.42	.43	.41	.42	.47	.41	.30	.45	.42	.42	.43	.39	.46	.43	.34	.41	.51
Fluency	.41	.39	.47	.40	.43	.38	.38	.42	.37	.31	.49	.47	.40	.41	.39	.46	.44	.34	.41	.50
TWE	.37	.36	.43	.33	.30	.28	.41	.35	.36	.29	.40	.40	.37	.36	.32	.44	.39	.30	.36	.44

R = Mean Correlation

R^d = Mean Disattenuated Correlation

TABLE 10

PAIRWISE CORRELATIONS OF TOEFL, TSE, AND TWE SUBTEST AND TOTAL SCORES
With 18 WRITTEN COMMUNICATION VARIABLES AVERAGED ACROSS 2 LEVELS OF REGISTER
AND ACROSS 6 PRAGMATIC FUNCTIONS
(N = 79 TOEFL, 71 TSE, 79 TWE)

	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16	W17	W18	R	R^d
TOEFL Total	.29	.34	.37	.37	.35	.38	.36	.38	.37	.35	.34	.37	.12	.12	.40	.34	.29	.13	.32	.42
Listening	.30	.37	.39	.42	.40	.41	.41	.42	.40	.36	.36	.41	.07	.18	.43	.36	.33	.17	.34	.47
Listening A	.23	.32	.33	.36	.37	.37	.38	.41	.37	.35	.35	.38	.05	.13	.36	.35	.30	.09	.31	.48
Listening B	.25	.32	.35	.38	.36	.35	.36	.36	.35	.26	.31	.39	.06	.18	.37	.27	.32	.15	.30	.49
Listening C	.32	.34	.34	.35	.32	.34	.34	.33	.32	.31	.28	.28	.06	.17	.40	.32	.24	.23	.30	.49
Structure and Written Expression	.20	.19	.22	.23	.21	.24	.19	.24	.24	.19	.20	.23	.04	.06	.26	.20	.16	.08	.19	.28
Structure	.23	.16	.21	.19	.18	.23	.16	.22	.20	.17	.18	.22	.12	.11	.20	.15	.18	.14	.18	.29
Written Expression	.14	.17	.18	.21	.18	.20	.17	.20	.22	.17	.17	.18	.00	.02	.25	.19	.12	.02	.16	.25
Vocabulary and Reading	.25	.31	.32	.30	.29	.33	.30	.31	.30	.35	.32	.31	-.02	.08	.33	.33	.25	.08	.26	.36
Vocabulary	.13	.13	.15	.14	.13	.14	.13	.18	.16	.23	.20	.16	-.04	-.01	.21	.20	.11	.01	.13	.21
Reading	.30	.39	.40	.37	.36	.41	.37	.36	.36	.38	.35	.37	.03	.13	.36	.32	.29	.11	.31	.44
TSE	.22	.32	.34	.34	.28	.34	.34	.30	.32	.24	.27	.31	-.04	.00	.29	.21	.31	.12	.25	.34
Comprehensibility	.27	.34	.36	.36	.34	.36	.34	.31	.33	.27	.30	.34	-.06	.02	.34	.26	.32	.20	.28	.39
Pronunciation	.37	.41	.40	.37	.38	.38	.37	.39	.40	.29	.27	.30	.05	.08	.33	.26	.31	.19	.31	.44
Grammar	.15	.27	.30	.26	.27	.28	.28	.25	.26	.19	.20	.24	-.04	.02	.24	.17	.26	.07	.20	.28
Fluency																				
TWE	.28	.38	.38	.34	.33	.37	.35	.38	.37	.34	.34	.34	.12	.10	.40	.39	.37	.39	.33	.45

R = Mean Correlation

R^d = Mean Disattenuated Correlation

BEST COPY AVAILABLE

TABLE 11

MULTIPLE REGRESSION ANALYSES WITH ORAL COMMUNICATIVE COMPETENCE AS THE DEPENDENT VARIABLE AND WITH ORAL COMMUNICATION VARIABLES BY COMPETENCE DOMAIN AS INDEPENDENT VARIABLES
(N = 948)

Competence Domain	Independent Variable		Coefficient	SE	t	Adj. R^2
Linguistic		Constant	.001	.017		.761*
	01	Pronunciation	.209	.112	1.86	
	02	Lexical Accuracy	-.403	.147	-2.75*	
	03	Lexical Complexity	1.051	.149	7.05*	
	04	Morphology	.146	.118	1.23	
	05	Syntactic Accuracy	.268	.139	1.92	
	06	Syntactic Complexity	.699	.141	4.95*	
Discourse		Constant	.001	.017		.756*
	07	Cohesion	1.162	.127	9.16*	
	08	Cohesence	.798	.127	6.28*	
Sociolinguistic		Constant	.000	.017		.752*
	09	Register Appropriacy	1.590	.095	16.67*	
	010	Cultural Appropriacy	.498	.094	5.28*	
Strategic		Constant	.000	.012		.866*
	011	Rate of Production	.405	.090	4.50*	
	012	Information Density	.552	.086	6.41*	
	013	Hesitation Phenomena	.620	.069	8.95*	
	014	Nonverbal Compensation	.166	.082	2.02**	
	015	Verbal Compensation	.244	.081	3.03*	

*p < .01 (2 tail)
**p < .05 (2 tail)
Oral communicative competence was obtained as an unweighted combination of oral strategic success (017) and listening comprehension (016); thus, both productive and receptive abilities were included.

TABLE 12

MULTIPLE REGRESSION ANALYSES WITH WRITTEN COMMUNICATIVE COMPETENCE# AS THE DEPENDENT VARIABLE AND WITH WRITTEN COMMUNICATION VARIABLES BY COMPETENCE DOMAIN AS INDEPENDENT VARIABLES (N = 948)

Competence Domain	Variable	Regression Coefficient	SE	t	Adj. R^2	Sel. R^2
Linguistic	Constant	3.123	.308		.228*	.343*
W1	Orthography	.167	.094	1.78		
W2	Lexical Accuracy	-.243	.183	-1.33		
W3	Lexical Complexity	.818	.177	4.62*		
W4	Morphology	.392	.123	3.18*		
W5	Syntactic Accuracy	-.230	.114	-2.02**		
W6	Syntactic Complexity	.372	.159	2.34**		
W13	Spelling	-.039	.037	-1.05		
W14	Punctuation	-.054	.064	-0.84		
Discourse	Constant	2.426	.230		.241*	.347*
W7	Cohesion	.912	.189	4.82*		
W8	Cohesence	.439	.186	2.36**		
Socio-linguistic	Constant	1.083	.274		.287*	.402*
W9	Register Appropriacy	.579	.157	3.68*		
W10	Cultural Appropriacy	1.198	.173	6.92*		
Strategic	Constant	.786	.247		.374*	.490*
W11	Rate of Production	.548	.136	4.02*		
W12	Information Density	.581	.110	5.30*		
W15	Verbal Compensation	.743	.086	8.64*		

*p < .01 (2 tail)
**p < .05 (2 tail)
Sel.R^2 = Squared multiple correlation for six highest ranking tasks
Written communicative competence was obtained as an unweighted combination of written strategic success (W17) and reading comprehension (W16); thus, both productive and receptive abilities were included.

TABLE 13

ORAL TASK EFFECTIVENESS IN REPRESENTING ELEMENTS OF COMMUNICATIVE COMPETENCE: DOMAIN ESTIMATES CORRELATED WITH RATED ORAL COMMUNICATIVE COMPETENCE* (STRATEGIC SUCCESS AND LISTENING COMPREHENSION) (N = 948)

TASK	LING.	DISC.	SOCIO.	STRAT.	MEAN R	RANK
1A Informal Responding	.837	.855	.810	.896	.850	9
1B Formal Responding	.820	.756	.794	.936	.827	11
2A Informal Requesting	.791	.795	.800	.907	.823	12
2B Formal Requesting	.897	.892	.903	.920	.903	5
3A Informal Persuading	.885	.878	.894	.931	.897	6
3B Formal Persuading	.896	.898	.905	.940	.910	4
4A Informal Complaining	.815	.813	.842	.922	.848	10
4B Formal Complaining	.897	.867	.878	.917	.890	7
5A Informal Apologizing	.951	.960	.968	.975	.964	2
5B Formal Apologizing	.929	.930	.936	.969	.941	3
6A Informal Summarizing	.874	.878	.877	.889	.880	8
6B Formal Summarizing	.975	.985	.986	.981	.982	1

*Oral communicative competence was obtained as an unweighted combination of oral strategic success (017) and listening comprehension (016); thus, both productive and receptive abilities were included.

TABLE 14

WRITTEN TASK EFFECTIVENESS IN REPRESENTING ELEMENTS OF COMMUNICATIVE COMPETENCE: DOMAIN ESTIMATES CORRELATED WITH RATED WRITTEN COMMUNICATIVE COMPETENCE* (STRATEGIC SUCCESS AND READING COMPREHENSION) (N = 948)

TASK	LING.	DISC.	SOCIO.	STRAT.	MEAN R	RANK
1A Informal Responding	.722	.741	.734	.837	.759	1
1B Formal Responding	.641	.635	.640	.696	.653	3
2A Informal Requesting	.511	.438	.549	.537	.509	5
2B Formal Requesting	.495	.521	.379	.584	.495	6
3A Informal Persuading	.369	.349	.411	.435	.391	10
3B Formal Persuading	.358	.419	.385	.466	.407	8
4A Informal Complaining	.570	.560	.560	.525	.554	4
4B Formal Complaining	.231	.183	.227	.215	.214	12
5A Informal Apologizing	.441	.437	.479	.561	.480	7
5B Formal Apologizing	.314	.350	.362	.459	.371	11
6A Informal Summarizing	.629	.658	.707	.727	.680	2
6B Formal Summarizing	.341	.390	.361	.535	.407	9

*Written communicative competence was obtained as an unweighted combination of written strategic success (W17) and reading comprehension (W16); thus, both productive and receptive abilities are included.

TABLE 15

CORRELATIONS OF ESTIMATED ORAL COMMUNICATIVE COMPETENCE VARIABLES WITH TOEFL, TSE, AND TWE TESTS AND SUBTESTS FOR SIX SELECTED ORAL COMMUNICATION TASKS (N = 426)

TEST	LING.	DISC.	SOCIO.	STRAT.	ORAL CC
TOEFL Total	.300	.294	.302	.304	.387
	(.335)	(.378)	(.338)	(.328)	(.308)
Listening	.265	.262	.270	.274	.259
	(.309)	(.306	(.316)	(.310)	(.291)
Listening A	.242	.237	.241	.244	.231
	(.313)	(.307)	(.313)	(.306)	(.287)
Listening B	.220	.215	.220	.217	.211
	(.304)	(.297)	(.305)	(.290	(.280)
Listening C	.225	.227	.240	.250	.230
	(.310)	(.313)	(.331)	(.333)	(.305)
Structure and Written Expression	.224	.221	.227	.221	.207
	(.274)	(.271)	(.279)	(.262)	(.244)
Structure	.322	.321	.318	.316	.300
	(.437)	(.436)	(.433)	(.415)	(.392)
Written Expression	.112	.109	.119	.112	.103
	(.147)	(.143)	(.157)	(.142)	(.130)
Vocabulary and Reading	.274	.265	.272	.277	.263
	(.320)	(.310)	(.318)	(.313)	(.295)
Vocabulary	.226	.227	.235	.237	.215
	(.299)	(.301)	(.313)	(.303)	(.274)
Reading	.263	.247	.253	.259	.254
	(.314)	(.295)	(.303)	(.299)	(.292)
TSE					
Comprehensibility	.374	.360	.365	.381	.367
	(.421)	(.406)	(.413)	(.416)	(.398)
Pronunciation	.340	.328	.334	.347	.337
	(.392)	(.379)	(.387)	(.388)	(.374)
Grammar	.369	.361	.369	.377	.365
	(.439)	(.430)	(.440)	(.434)	(.418)
Fluency	.362	.350	.357	.376	.360
	(.423)	(.410)	(.419)	(.426)	(.405)
TWE	.311	.318	.321	.327	.314
	(.359)	(.368)	(.372)	(.365)	(.349)

.098 = critical value for $p < .05$ (one tail)
.128 = critical value for $p < .01$ (one tail)
Note: Figures in parentheses are disattenuated correlations.
ORAL CC = Unweighted combination of oral strategic success (017) and listening comprehension (016).

TABLE 16

CORRELATIONS OF ESTIMATED WRITTEN COMMUNICATIVE COMPETENCE VARIABLES WITH TOEFL, TSE, AND TWE TESTS AND SUBTESTS FOR SIX SELECTED WRITTEN COMMUNICATION TASKS (N = 426)

TEST	LING.	DISC.	SOCIO.	STRAT.	WRITTEN CC
TOEFL Total	.304	.289	.277	.324	.318
	(.456)	(.429)	(.394)	(.431)	(.353)
Listening	.347	.351	.316	.360	.356
	(.545)	(.546)	(.470)	(.501)	(.414)
Listening A	.317	.324	.291	.334	.331
	(.552)	(.558)	(.480)	(.515)	(.426)
Listening B	.282	.286	.246	.294	.304
	(.524)	(.526)	(.433)	(.484)	(.418)
Listening C	.299	.299	.281	.305	.287
	(.554)	(.548)	(.493)	(.501)	(.393)
Structure and Written Expresion	.173	.156	.150	.214	.192
	(.285)	(.255)	(.234)	(.313)	(.234)
Structure	.177	.162	.157	.205	.186
	(.323)	(.293)	(.272)	(.332)	(.251)
Written Expression	.133	.118	.113	.173	.154
	(.235)	(.206)	(.189)	(.271)	(.201)
Vocabulary and Reading	.243	.218	.227	.248	.253
	(.381)	(.339)	(.338)	(.345)	(.294)
Vocabulary	.112	.112	.141	.157	.169
	(.199)	(.197)	(.238)	(.248)	(.222)
Reading	.299	.260	.252	.272	.272
	(.480)	(.413)	(.384)	(.387)	(.323)
TSE					
Comprehensibility	.263	.263	.241	.279	.296
	(.399)	(.395)	(.346)	(.375)	(.332)
Pronunciation	.289	.276	.267	.304	.302
	(.449)	(.424)	(.393)	(.419)	(.347)
Grammar	.293	.283	.251	.268	.268
	(.469)	(.448)	(.381)	(.380)	(.317)
Fluency	.228	.243	.216	.236	.261
	(.359)	(.379)	(.322)	(.329)	(.304)
TWE	.276	.254	.236	.268	.304
	(.429)	(.390)	(.347)	(.369)	(.349)

.098 = critical value for $p < .05$ (one tail)
.128 = critical value for $p < .01$ (one tail)
Note: Figures in parentheses are disattenuated correlations.
WRITTEN CC = Unweighted combination of written strategic success (W17) and reading comprehension (W16).

APPENDIX

COMMUNICATIVE ELICITATION TASKS: I. Aural/Oral

Strategic Functions

(Minutes)

1. Responding to requests for academic information

(0-5) a) [Informal] Hello, I'm _____. What's your name? What country are you from? What's your academic specialization/interest? Tell me how you got interested in this particular specialization.

(6-10) b) [Formal] Imagine I am a university admissions officer and I am interviewing you for admission. Tell how many months you have studied English. Tell how many months you have lived in an English-speaking country. Describe in detail your academic background beginning with elementary school.

2. Requesting academic information

(11-15) a) [Informal] Ask me as a friend to tell you about my educational background. Keep the conversation going.

(16-20) b) [Formal] Imagine I am your professor and you are coming to see me to ask about your progress in my course. Introduce yourself and request this information. Try to get a lot of information.

3. Persuading change

(21-25) a) [Informal] Imagine I am your friend who decided not to go to university. Persuade me to attend university.

(26-30) b) [Formal] Imagine I am your professor and you want me to raise your grade in English class. Persuade me to change your grade.

4. Complaining

(31-35) a) [Informal] Imagine I am your roommate, and you want to tell me about some course you disliked in school. Tell me why it was bad.

(36-40) b) [Formal] Imagine I am your professor and you want to tell me that my course was not good enough. Introduce yourself and tell me about my course.

5. Apologizing for inadequate behavior

(41-45) a) [Informal] Imagine I am your friend who invited you to a party, but now you can't go because you have an examination at that time. Apologize for not coming.

(46-50) b) [Formal] Imagine that I am your professor and you are late in turning in a term paper. Introduce yourself, tell me about the problem and apologize for being late.

6. Summarizing information received

(51-55) a) [Informal] Listen to the following information and then summarize it for me in your own words. (Read aloud the following passage:)

Letter Home

Dear Dad and Mom,

It is now only one month since I left home to come here to school. I am having a wonderful time, but I miss you both very much. I wish you could meet my teachers and see the classes I attend. I am learning a

lot, but it has not always been easy. Some of my classes require a lot of preparation, so I spend a lot of time in the evenings studying.

Last week I made a friend from China. She is studying English in the same class with me. I also meet people from many other countries. Part of my education is no doubt getting to meet people from so many different cultures. I am also getting used to different kinds of food.

I've been very well and I still have enough money to meet all my expenses here. I hope everything is well with you. Please write and share all the news from home.

Love,
Maria

(56-60) b) [Formal] Listen to the following information and then summarize all of the important points as you would for a class in school. (Read aloud the following passage:)

Left Out in the Cold

When the UN Decade of Disabled Persons was launched in 1983, specialists estimated that the world had been spending only 1 cent on each handicapped man, woman and child annually. To improve their lot would have required a modest increase to 25 cents, the specialists said.

Five years and a timid economic recovery later, most societies have advanced little, if at all, beyond the financial starting line. Meanwhile, the number of people around the world who are physically or mentally disabled--victims of war, accident, malnutrition and disease--has surpassed the 500 million mark. Their predicament is particularly grave in developing countries where facilities are woefully inadequate or nonexistent. Many turn to begging as a means of livelihood.

Available funds could be used more effectively if they are channeled to programs that promote self-employment of disabled persons. This approach has become a major feature of vocational rehabilitation projects. Group employment can also be successful, as shown by a project in Zimbabwe that helped some 100 handicapped people make a living from the cultivation and processing of sunflowers.

(In all of the above activities, probe as necessary to keep the conversation going. For example, ask, "Why do you think that? Why do you feel that way? Anything else?")

COMMUNICATIVE ELICITATION TASKS: II. Reading/Writing

Strategic Functions

(Minutes)

1. Responding to requests for academic information

(0-5) a) [Informal] Write your name, the name of the country where you grew up, and a paragraph or two about a memorable experience you or someone you know had in school.

(6-10) b) [Formal] Write a letter applying to a university graduate school in your subject area. Tell how your academic background qualifies you for admission.

(11-15) a) [Informal] Write a note asking a friend to help you get ready for an English test. Ask about the best ways to study.
(16-20) b) [Formal] Write a letter to a college asking for application procedures, tuition and housing costs, and other important information about the college.

3. Persuading change

(21-25) a) [Informal] Write a paragraph or two trying to convince a friend that he/she should come and attend the same school you are attending.
(26-30) b) [Formal] Write a letter to the governor of a state persuading him/her to make university education free of charge to all who are qualified and want it.

4. Complaining

(31-35) a) [Informal] Write a note to a neighbor asking him/her to keep the noise level down in the evenings so you can study.
(36-40) b) [Formal] Write a letter to a professor explaining that he/she is asking for too much homework for the course you are taking.

5. Apologizing for inadequate behavior

(41-45) a) [Informal] Write a note to your friend explaining that you are sorry but you cannot keep your promise to visit that friend's home during the holidays.
(46-50) b) [Formal] Write a note to your English professor explaining that you will be one week late in completing a term paper.

6. Summarizing information received

(51-55) a) [Informal] Read and summarize in your own words the ideas in the following dialogue:

Students Prepare for Final Examinations

Maria: I don't know about you, Ahmad, but I've tried everything I know to get a good grade in English class, but I'm still having trouble.
Ahmad: What seems to be the problem, Maria?
Maria: Well, my vocabulary is very limited and I still have difficulty pronouncing some English words correctly.

Ahmad: I used to have a lot of trouble with vocabulary too, but then I began reading interesting short stories in magazines like Reader's Digest. I found that fun and interesting, and it helped me a lot. I also listen to the news on TV and try to repeat sentences the way they're said by the newscasters. That helps my pronunciation and fluency a lot. My English is still not perfect, but I've come a long way since the beginning of the term.
Maria: Thanks for those ideas, Ahmad. I'll try them out as I get ready for the final exam next week.
Ahmad: Good luck, Maria!

(56-60) b) [Formal] Read and summarize in brief report form the information in the following passage:

Marine Scientists Join Search for Greenhouse Accord

GENEVA--Leading marine research and training institutions in most coastal countries are to participate in an elaborate global study to assess the likely impact of climatic change due to the "greenhouse effect" on sensitive sectors of the environment.

The study, launched in search of international accord on a collective, rational response to rising sea levels and global warming caused by pollutant gases in the atmosphere, is coordinated by the United Nations Environment Programme (UNEP). This is the first big response of the UN to the problem since it was stated by scientists from 48 countries at a recent Toronto conference on the changing atmosphere.

Scientists describe the projected "Greenhouse" changes in the climate as the biggest long-term threat to humanity short of nuclear war; and they have called for a significant reduction in carbon dioxide emissions. The UNEP study is intended to help governments decide on what actions to take in order to limit the extent of the damage.

A special UNEP scientific task force has just been established to identify and address the likely impact of climactic change on important areas such as public health, water resources, marine and terrestrial ecosystems and coastal regions. Other international agencies involved in the program include the World Meteorological Organization and the International Council of Scientific Unions.

References

Aitkin, M. & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. The Journal of the Royal Statistical Society: Series A (General), 149, 1-43.

Bachman, L. (1986). The Test of English as a Foreign Language as a measure of communicative competence. In C. W. Stansfield (Ed.), Toward communicative competence testing: Proceedings of the second TOEFL invitational conference (TOEFL Research Report No. 21). Princeton, NJ: Educational Testing Service.

Bachman, L. F., (1990). Fundamental considerations in language testing. Oxford: Oxford University Press.

Bachman, L. F., & Clark, J. L. D. (1987). The measurement of foreign/second language proficiency. Annals, AAPSS, 490, 20-33.

Bachman, L. F., & Palmer, A. (1981). The construct validation of the FSI oral interview. Language Learning, 31(1), 67-86.

Bachman, L. F. (1990). Fundamental Considerations in Language Testing. Oxford: Oxford University Press.

Brennan, R. L. (1991). Statistical models for behavioral observations: A review. Journal of Educational Statistics, 16 (3), 253-266.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1, 1-47.

Candlin, C. N. (1986). Explaining communicative competence: Limits of testability? In C. W. Stansfield (Ed.), Toward communicative competence testing: Proceedings of the second TOEFL invitational conference (TOEFL Research Report No. 21). Princeton, NJ: Educational Testing Service.

Carroll, J. B. (1961). Fundamental considerations in testing English proficiency of foreign students. In, Testing the English proficiency of foreign students. Washington, DC: Center for Applied Linguistics.

Carroll, B. C. (1980). Testing communicative performance: An interim study. Oxford: Pergamon Press.

Duran, R. P., Canale, M., Penfield, J., Stansfield, C. W., & Liskin-Gasparro, J. E. (1985). TOEFL from a communicative viewpoint on language proficiency: A working paper (TOEFL Research Report No. 17). Princeton, NJ: Educational Testing Service.

Faerch, C., & Kasper, G. (1984). Two ways of defining communication strategies. Language Learning, 34 (1), 45-63.

Halliday, M. A. K. (1973). Relevant models of language. In M. A. K. Halliday (Ed.), Explorations in the functions of language. New York: Elsevier North-Holland.

Halliday, M. A. K. (1976). The form of functional grammar. In G. Kress (Ed.), Halliday: System and function in language. Oxford: Oxford University Press.

Hatch, E., & Long, M. (1980). Discourse analysis, what's that? In D. Larsen-Freeman (Ed.), Discourse analysis in second language research. Rowley, Mass.: Newbury House.

Hymes, D. (1972). On communicative competence. In B.J. Pride and J. Holmes, (Eds.), Sociolinguistics. Harmondsworth: Penguin.

Joos, M. (1967). The five clocks. New York: Harcourt, Brace, Jovanovich.

Lado, R. (1961). Language testing. New York: McGraw-Hill.

Messick, S. (1987). Validity (Research Report RR-87-40). Princeton, NJ: Educational Testing Service. Also in R. L. Linn (Ed.), Educational Measurement (3rd ed.). New York: MacMillan, 1988.

Oller, J. W., Jr. (1979). Language tests at school. London: Longman.

Oller, J. W., Jr. (1986). Communication theory and testing: What and how. In C. W. Stansfield (Ed.), Toward communicative competence testing: Proceedings of the second TOEFL invitational conference (TOEFL Research Report No. 21). Princeton, NJ: Educational Testing Service.

Savignon, S. J. (1986). The meaning of communicative competence in relation to the TOEFL program. In C. W. Stansfield (Ed.), Toward communicative competence testing: Proceedings of the second TOEFL invitational conference (TOEFL Research Report 21). Princeton, NJ: Educational Testing Service.

Spolsky, B. (1985). What does it mean to know how to use a language? An essay on the theoretical basis of language testing. Language Testing, 2(2), 180-191.

Spolsky, B. (1989). Conditions for Second Language Learning. Oxford: Oxford University Press.

Stansfield, C. W. (1983). Review of recent TOEFL program research and directions for the future. A paper submitted to the TOEFL Research Committee and to the TOEFL Policy Committee. Princeton, NJ: Educational Testing Service.

Stansfield, C. W. (Ed.) (1986). Toward communicative competence testing: Proceedings of the second TOEFL invitational conference (TOEFL Research Report No. 21). Princeton, NJ: Educational Testing Service.

Wright, B. D., & Masters G. N. (1982). Rating scale analysis: Rasch measurement. Chicago: MESA Press.

Printed on Recycled Paper

57906-01198 • Y121M.5 • 275578 • Printed in U.S.A.